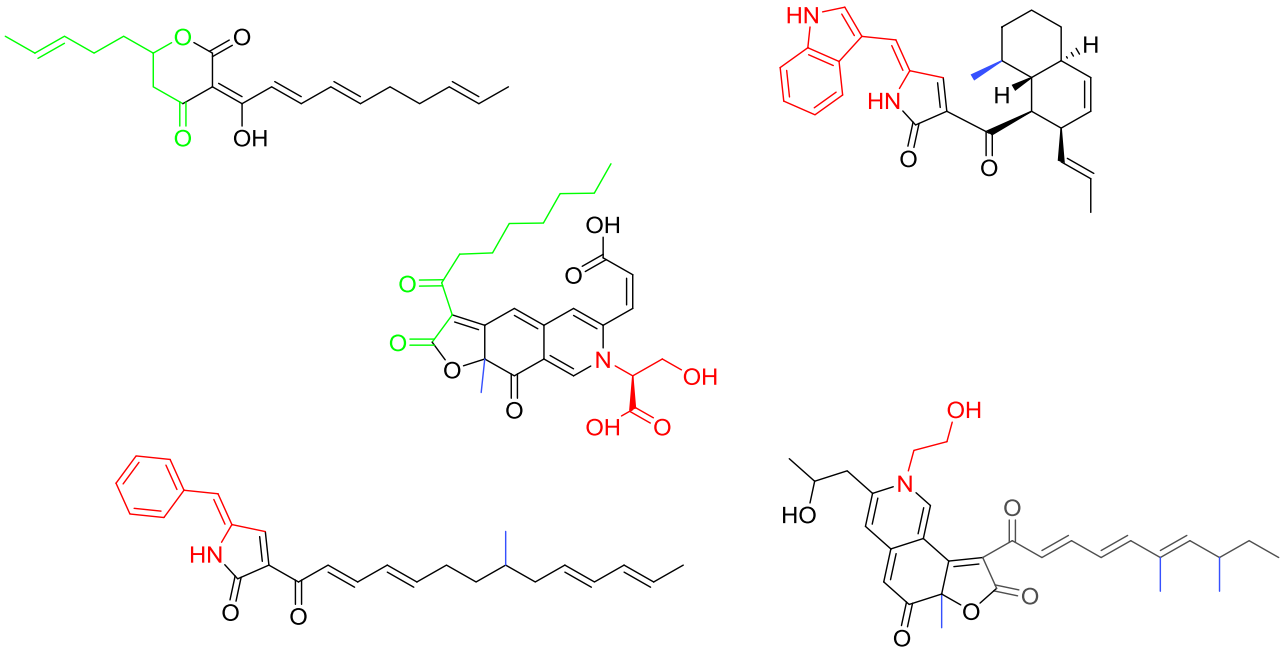
Discovery and Characterization of Fungal Natural Product Biosynthetic Pathways

Isbrandt, Thomas

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Discovery and Characterization of Fungal Natural Product Biosynthetic Pathways



PhD Thesis

Thomas Isbrandt

Department of Biotechnology and Biomedicine

Technical University of Denmark

February 2018

Preface

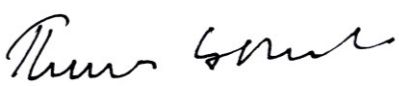
The present thesis is the result of a little over three years of PhD studies in the Natural Product Discovery group at the Department for Biotechnology and Biomedicine (DTU Bioengineering) at the Technical University of Denmark (DTU). The project started in December 2014 and finished in February 2018 and was funded by DTU.

During the project, I have been involved in many different kinds of research within the field of natural product chemistry, and I have had several collaborators who all deserve a big *thank you*. Firstly, I would like to thank my supervisor Thomas Ostenfeld Larsen for excellent guidance throughout the project, and for always having a moment to spare for discussion.

Secondly, I want to thank Maria Lund Nielsen for a great collaboration that forced me to learn about gene clusters and molecular biology; Gerit Tolborg and Anders Ødum for our endeavours into the world of fungal pigments; and Jakob Blæsbjerg Hoof for indispensable help and work on creating mutants for me to analyse. In addition a big *thank you* goes to Kasper Enemark-Rasmussen and Charlotte Held Gotfredsen for running all of my NMR samples.

Thirdly, I would like to give a special thanks to all the people in 221, who contributed to making a brilliant environment, both academically and socially. Especially, I would like to thank Chris Phippen for fruitful discussions about all things chemistry and non-chemistry, and Karolina Subko for always supplying the office with chocolate and other goodies. Furthermore, I would like to thank Magnus Hallas-Møller, Andreas Heidemann, Daniel Killerup Svenssen, Aaron Andersen, Alexander Rosenkjær, Sara Kildgaard, Silas Anslem Rasmussen, Anja Madsen, and Sietske Grijseels, as well as all the people in 223.

Lastly, I want to thank my family and friends for support during the last three years, especially my bois in PIQAIA for giving me some much needed breaks from all the science stuff!



Thomas Isbrandt

Summary

The kingdom of fungi encompasses an enormous amount of species, each of which have a unique ability to produce enzymes and secondary metabolites. The secondary metabolites fill a variety of roles in the development and life of a fungus, and many secondary metabolites also possess various bioactivities or other properties of relevance to humans, such as antibacterial, antifungal, anticancer, neuroactive, and toxic properties. Additionally, some species produce compounds with industrial relevance, such as pigments, flavouring agents or other food additives. As a result of a great decrease in sequencing costs during the last decade, the doors to new bioinformatics approaches within natural product discovery have been opened. Similarly, advances in analytical equipment has made more in-depth and accurate analyses possible, further exposing the potential of the fungi which are being analysed. Combining these techniques with molecular biology have made it possible to more confidently link secondary metabolites to their responsible genes, suggest biosynthetic pathways, and undertake metabolic engineering with the goal of expanding the secondary metabolite catalogue. In this project natural product discovery has been approached from several different ways, including spectroscopy and spectrometry guided discovery, as well as genetics based discovery and engineering.

Spectroscopy-guided analysis of the filamentous fungus *Talaromyces atroroseus* lead to the discovery of a novel red azaphilone pigment (Appendix 1 and 2). Further investigation of the chemical potential by dereplication revealed that in fact a whole range of red pigments were produced by the fungus, differing by the incorporation of various amino acids, all structurally elucidated with one- and two-dimensional nuclear magnetic resonance spectroscopy. These compounds were named atrorosins, and were found to be a new class of *Monascus* pigments. By cultivating the fungus under controlled conditions in bioreactors, it was possible to design a fermentation process in which the identity of the incorporated amino acid could be decided, while simultaneously achieving product yields in the 'gram per liter'-scale. Furthermore, three novel yellow and violet azaphilones from *A. neoglaber* was characterised (Appendix 4), by nuclear magnetic resonance spectroscopy, and high resolution mass spectrometry in conjunction with deuterium labelling, could be used to further confirm the structure of one of these compounds.

Several tools has been developed for genetic engineering of microorganisms, with CRISPR/cas9 being one of the most famous ones in the last few years. This technique can be used to alter the products or elucidate the route of a given biosynthetic pathway, or to increase the yields in a bioprocess (Appendix 6).

Aspergillus brasiliensis is a filamentous fungus belonging to the black Aspergilli, section *Nigri*, also dealt with throughout this project (Appendix 5). Dereplication of the secondary metabolite profile of this relatively newly described species resulted in identification of a range of unknown and possibly novel compounds. Of particular

interest was one compound, produced in large amount and possessing a unique absorption spectra. Upon purification and structural elucidation the compound was found to indeed be a previously unknown polyketide-fatty acid hybrid, that was named brasenol A1. Bioactivity testing showed brasenol A1 to have mild antibacterial properties, and investigation of other fungi in the section *Nigri* revealed that also *A. carbonarius* was also a producer, of brasenol and several putative analogues. Using comparative bioinformatics and gene deletions, the responsible gene cluster was identified, and turned out to encode a highly reducing polyketide synthase, BrsA, a hydrolase, BrsB, and an esterase, BrsC. When heterologously expressing the genes *brsA* and *brsB* together in three different fungi (*A. nidulans*, *A. sydowii* and *A. oryzae*), brasenols were produced in vast amounts, and characterisation of three analogues was possible, brasenol B1, B2 and C1. However, even more interesting, sporulation was interrupted in all three hosts, possibly as a result of disruption of the fatty acid/oxylipin metabolism in the fungi.

The potential of bioinformatics is by no means limited to elucidation of biosynthetic pathways. Engineering of gene clusters in order to make synthetic natural products is another approach which have been explored in this project (Appendix 7). With the aim of engineering novel cytochalasin analogues, the polyketide synthase-non ribosomal synthetase (PKS-NRPS) responsible for biosynthesis of cytochalasin E in *Aspergillus clavatus*, CcsA, and its native trans-acting enoyl reductase, CcsC, was heterologously expressed in *A. nidulans*. This led to production of an unknown secondary metabolite, which was named niduclavin. Structural characterisation showed the compound to consist of an tri-methylated octaketide linked to a phenylalanine moiety through a five-membered lactam. The compound was furthermore found to contain a decalin system rather than the tricyclic isoindolone system normally observed for cytochalasins, and in fact be more similar to compounds such as talaroconvolutin A, myceliothermophin E, and equisetin. Bioinformatics led to identification of a homologue of the *A. clavatus* PKS-NRPS encoding gene, *ccsA*, in the rice blast fungus *Magnaporthe oryzae*, and this gene, *syn2*, along with its native trans-acting enoyl reductase, *rap2*, was similarly heterologously expressed in *A. nidulans*. The natural product from the *syn2* gene had not previously been identified, but heterologous expression in *A. nidulans* led to production of a compound similar to niduclavin, consisting of an singly methylated octaketide linked to a tryptophan moiety through a lactam. The compound was named niduporthin. Common for both niduclavin and magnaporthin was addition of a double bond in the α/β-position of the amino acid, a modification most likely caused by native enzymes in *A. nidulans*. In order to obtain synthetic analogues of the compounds obtained from expression of the two PKS-NRPS genes, the PKS and NRPS modules were swapped between the two hybrid genes. This resulted in two chimeric analogues, niduchimaeralin A and B. Based on tandem MS experiments, niduchimaeralin A and B were determined to indeed be the expected swapped versions of niduclavin and niduporthin.

In a different approach to genetic engineering, the lovastatin producing PKS, LovB, was heterologously expressed in *A. nidulans*, along with the NRPS module of the PKS-NRPS CcsA, from *A. clavatus* (Appendix 8). LovB is a well-studied PKS, also encoding a condensation domain, usually only found in NRPS or hybrid PKS-NRPS genes. In order to investigate the PKS and the role of the C domain, two synthetic hybrid versions of a LovB/ccsA PKS-NRPS was made. Mutant 1 consisted of the LovB PKS and C domain linked to the CcsA NPRS module lacking the C domain, and Mutant 2 consisted of only the LovB PKS, without the C domain, linked to the whole CcsA NRPS module. Mutant 1 was found to produce dihydromonacolin L, a lovastatin precursor, whereas NMR structural elucidation revealed Mutant 2 to produce a novel PK-NRP hybrid, named terreclavin, constructed from phenylalanine and a linear octaketide linked via the same five membered lactam as seen in the CcsA and Syn2 compounds. Lovastatin is a nonaketide, and various reasons could be the cause of the shorter polyketide chain observed in terreclavin. Speculations about the exact role of the LovB C domain has previously been proposed, and while functions such as a Diels-Alderase-like activity has been proposed, this study strengthens the hypothesis that the C domain is indeed crucial for the biosynthesis of lovastatin in *A. terreus*.

Altogether, this project has dealt with several aspects of natural product discovery, from spectroscopy guided discovery and dereplication, to bioinformatics and molecular biology-based approaches to uncover new compounds with potentially useful applications. Discovery of a new azaphilone, showed that using dereplication, it was possible to characterise a whole new class of *Monascus* pigments for potential use as natural food additives. Furthermore, the use of high resolution mass spectrometry during structural characterisation proved a useful tool for both structural elucidation and confirmation. Through collaborations with molecular biologists, characterisation of novel PKS-NRPS products and elucidation of the biosynthetic pathway of a novel biomarker was done. In conclusion, the possibilities within the field of natural product discovery are still great, and with the development of increasingly advanced analytical tools, and methods within bioinformatics and molecular biology, the prospects for discovering and engineering of novel and useful molecules are as promising as ever.

Dansk resumé

Svamperiget omfatter en enorm mængde arter, som hver især er i stand til at producere unikke enzymer og sekundære metabolitter. De sekundære metabolitter er ansvarlige for et væld af funktioner i forbindelse med skimmelsvampes udvikling og livscyklus, og mange sekundære metabolitter er i besiddelse af diverse aktiviteter som kan udnyttes af mennesker, som for eksempel, antibakterielle, antifungale, anticancer, neruroaktive, og toksiske egenskaber. Derudover er nogle arter i stand til at lave stoffer med industriel relevans, så som pigmenter, aromastoffer eller andre tilsætningsstoffer til føde varer. Som følge af det markante fald i prisen for gensekventering i løbet af de sidste ti år, er dørene blevet åbnet til nye bioinformatiske fremgangsmåder inden for opdagelse af naturstoffer. Samtidig har udviklingen af instrumenter gjort kemiske analyser mere nøjagtige, hvorved det kemiske potentiale i svampe er blevet yderligere afsløret. Ved at kombinere disse teknikker med værktøjer indenfor molekylær biologi, er det blevet muligt at give gode bud på sammenhængen mellem sekundære metabolitter og de gener som er ansvarlige for deres produktion, og der er samtidig blevet lettere at arbejde med det genetiske materiale og derved yderligere udvide svampenes kemiske repertoire. I dette projekt er naturstofkemi blevet grebet an fra flere forskellige retninger, for eksempel med udgangspunkt i spektroskopi og spektrometri, men også på baggrund genetik.

Spektroskopi-baseret analyse af skimmelsvampen *Talaromyces atroroseus* førte til opdagelsen af et nyt rødt azaphilon-afledt pigment (Appendix 1 og 2). Yderligere undersøgelse af det kemiske potentiale ved hjælp af dereplikering, vise at svampen rent faktisk var i stand til at producere en række forskellige røde farstoffer, som hver især varierede ved at indeholde forskellige aminosyrer, afsløret ved hjælp af en- og todimensionel kernemagnetisk resonans spektroskopi. Stofferne blev navngivet atrorosiner, og viste sig at høre til en stofklasse kaldet *Monascus*-pigmenter. Ved at dyrke svampen under kontrollerede betingelser i bioreaktorer, var det muligt af udforme en fermenteringsprocess hvori indbygningen af aminosyre kunne kontrolleres og samtidig give udbytter i størrelsesordenen ”gram per liter”. Derforuden blev yderligere tre azaphiloner fra skimmelsvampen *Aspergillus neoglaber* isoleret (Appendix 4), og karakteriseret ved hjælp af NMR spektroskopi, og ved brug af deuterium-mærkning sammen med højopløst massespektrometri var det muligt at yderligere bekræfte strukturen af det ene af stofferne.

Der er udviklet mange værktøjer til genetisk manipulation af mikroorganismer, med CRISPR/cas9 værende et af de senere års mest berømte. Teknikken kan blandt andet bruges til at ændre produkterne af, eller afsløre hvorved en given biosyntetisk rute er sammensat, eller bruges til at optimere udbytterne ved en bioprocess (Appendix 6).

Den sorte aspergille *Aspergillus brasiliensis* er endnu en skimmelsvamp som er blevet arbejdet med i dette projekt (Appendix 5). Dereplikering af de sekundære metabolitter produceret at denne svamp resulterede i identifikation af flere ukendte og hidtil ubeskrevne stoffer. Særligt interessant var et stof, som blev produceret i særlig store mængder og havde et unikt absorptionsspektrum. Oprensning og strukturopklaring viste at stoffet var et hidtil ukendt polyketid og fedtsyre hybrid, som blev navngivet brasenol A1. Stoffet viste sig, til en vis grad, at have antibakterielle egenskaber, ved undersøgelse af andre nærtbeslægtede sorte aspergiller viste *A. carbonarius* sig også at producere stoffet samt flere formodede analoger. Ved brug af komparativ bioinformatik og gendeletioner, blev gen-clustret ansvarlig for brasenol-produktion identificeret, og det viste til at dette var sammensat af gener kodende for en reducerende polyketidsyntase, BrsA, en hydrolase, BrsB, og en esterase, BrsC. Heterolog udtryk af generne i tre forskellige svampe (*A. nidulans*, *A. sydowii*, og *A. oryzae*) resulterede til produktion af store mængder brasenol, hvorved karakterisering af yderligere tre analoge stoffer var mulig. Mere interessant var dog at sporedannelse i alle tre værter stoppede, muligvis som resultat af forstyrrelse af fedtsyremetabolismen i svampene.

Potentialet i bioinformatik er langt fra begrænset til afklaring af biosynteseveje. Manipulation af genclustre med henblik på at fremstille syntetiske naturstoffer er en anden tilgang som er blevet udnyttet i dette projekt (Appendix 7). Med henblik på produktion af nye cytochalasin-afledte stoffer, blev polyketidsyntase-non-ribosomal peptid syntetasen (PKS-NRPS) CcsA udtrykt sammen med enoyl reduktasen CcsC, i *Aspergillus nidulans*. Det førte til produktion af en ukendt sekundær matabolit, som blev navngivet niduclavin. Strukturopklaring viste at stoffet bestod af et trimethyleret oktaketid sammensat med aminosyren phenylalanin via en fem-leddet lactam. Stoffet viste sig yderligere at indeholde en decalin ring frem for det makrocykliske system som normal ses i cytochalasin, og det var således mere lig stoffer som talaroconvolutin A, myceliothermophin E og equisetin. Ved hjælp af bioinformatik lykkedes det at identificere en homolog til ccsA i den risinficerende svamp *Magnaporthe oryzae*, Syn2, som ligeledes blev heterologt udtrykt i *A. nidulans* sammen med den tilhørende enoyl reduktase Rap2. Det naturlige produkt fra *syn2*-genet var ikke tidligere beskrevet, men forsøgene i *A. nidulans* førte til produktion af et stof som mindede meget om niduclavin, med bestod af et enkelt-methyleret oktaketid sammensat med aminosyren tryptophan via samme fem-leddede lactam. Stoffet blev navngivet niduporthin. Fælles for både niduclavi og niduporthin, var at begge stoffer indeholdt en dobbeltbinding i α/β-positionen i aminosyredelen, en modifikation som sandsynligvis er et resultat af enzymer i *A. nidulans*. For at opnå produktion af syntetiske analoger af de to stoffer, blev PKS og NRPS modulerne i CcsA og Syn2 ombyttet. Det resulterede i produktion af to kimære stoffer, niduchimaeralin A og B og ved hjælp af tandem MS eksperimenter var det muligt af bekræfte identiteten af de to nye stoffer.

I et andet studie blev den lovastatin-producerende PKS LovB fra *A. terreus* udtrykt i *A. nidulans* sammen med NRPS modulet fra PKS-NRPS'en ccsA (Appendix 8). LovB er en velstuderet PKS som også indeholder et kondensationsdomæne (C domæne), typisk kun set i PKS-NRPS hybrider. For at undersøge PKS'en og hvilken rolle C domænet spiller, blev to udgaver af en LovB/CcsA hybrid designet. Mutant 1 bestod af LovB's PKS og C domæne sat sammen med NRPS modulet fra ccsA, uden dennes C domæne, og Mutant 2 bestod af LovB's PKS direkte sammensat med hele CcsA's NRPS modul. Ved kemisk analyse viste Mutant 1 sig at producere stoffet dihydromonacolin L, en forløber til lovastatin, hvorimod Mutant 2 på baggrund af NMR studier viste sig at producere et nyt PKS-NRPS stof, terreclavin, bestående af en phenylalanin-del sammensat med et lineært oktaketid via den samme type lactam som også var observeret i CcsA og Syn2 produkterne. Lovastatin er normalt et nonaketid og flere årsager kan være årsag til den kortere PK-del konstateret i terreclavin. Der har være spekulationer om C domænets rolle i lovasatin biosyntese, og mens funktioner som Diels-Alderase-lignende aktivitet har været blandt de foreslåede, styrker dette studie hypotesen om a C domænet uden tvivl er spiller en nøglerolle i lovastatin-biosyntese i *A. terreus*.

Dette Ph.d.-projekt har håndteret adskillige aspekter indenfor naturstofkemi, fra spektroskopibaseret opdagelse og dereplikering, til bioinformatik og molekylærbiologi-baserede tilgange til opdagelse af nye forbindelser med potentielt brugbare egenskaber. Opdagelsen af en ny azaphilon, viste at brug af dereplikering gjorde opdagelsen af en ny type *Monascus* pigmenter til potentiel brug i fødevarer mulig. Derforuden viste højopløst massespektrometri sig at være et særdeles brugbart redskab i opklaring og bekræftelse af nye strukturer, hvis brugt parallelt med NMR spektroskopi. Gennem samarbejde med molekylærbiologer, blev flere nye PKS-NRPS hybridstoffer karakteriseret, og biosyntesevejen for en ny biomarkør blev klarlagt. For at opsummere er mulighederne indenfor opdagelse af nye naturstoffer stadig mange, og med udviklingen af stadig mere avancerede analytiske redskaber, samt nye metoder indenfor bioinformatik og molekylærbiologi, ser udsigten til opdagelse af endnu flere nye naturstoffer i fremtiden lovende ud.

List of publications and conference contributions

Publications

Peer-reviewed

Linker Flexibility Facilitates Module Exchange in Fungal Hybrid PKS-NRPS Engineering. Nielsen, M. L., Isbrandt, T., Petersen, L. M., Mortensen, U. H., Andersen, M. R., Hoof, J. B., Larsen, T. O., PLoS One, vol 11, issue 8, 2016, DOI: 10.1371/journal.pone.0161199

Genes Linked to Production of Secondary Metabolites in Talaromyces atroroseus Revealed Using CRISPR-Cas9. Nielsen, M. L., Isbrandt, T., Rasmussen, K. B., Thrane, U., Hoof, J. B., Larsen, T. O., Mortensen, U. H., PLoS One, vol 12, issue 1, 2017, DOI: 10.1371/journal.pone.0169712

Establishing novel cell factories producing natural pigments in Europe. Tolborg, G., Isbrandt, T., Larsen, T. O., Workman, M. in *Bio-pigmentation and Biotechnological Implementations* (ed. Singh, O. V.) 23–51 (John Wilety & Sons, Inc., 2017), DOI: 10.1002/9781119166191.ch2

Submitted

Unique processes yielding pure azaphilones in Talaromyces atroroseus. Tolborg, G., Isbrandt, T., Ødum, A., Larsen, T. O. & Workman, M.

In preparation

Structure and genetic origin of novel class of polyketide biomarkers from Aspergillus brasiliensis, brasenols. Isbrandt, T., Hoof, J. B. & Larsen, T. O.

Atrorosins: a new subgroup of Monascus pigments from Talaromyces atroroseus. Isbrandt, T., Tolborg, G., Workman, M. & Larsen, T. O.

New azaphilones from Aspergillus neoglaber. Isbrandt, T. & Larsen, T. O.

Reviving the Extinct PKS-NRPS Activity of the Lovastatin Nonaketide Synthase LovB. Nielsen, M. L., Isbrandt, T., Mortensen, U. H., Hoof, J. B., Andersen, M. R., Larsen, T. O.

Biosynthesis of acurin A and B, two novel isomeric fusarin C-like compounds from Aspergillus aculeatus. Nielsen, M. L., Wolff, P. P., Petersen, L. M., Andersen, L. N., Isbrandt, T., Holm, D. K., Mortensen, U. H., Nødvig, C. S., Larsen, T. O., Hoof, J. B.

Conference contributions

Prediction of Secondary Metabolite Encoding Genes Based on Chemical Structure Analysis. Isbrandt, T., Hoof, J. B., Nielsen, M. L., Gezgin, Y., Nielsen, K. F., Larsen, T. O., Nordic Natural Product Conference 2017 , 2017, Odense, DK.

Discovery and structure elucidation of natural pigments from *Talaromyces atroroseus*. Kroll-Møller, P., Isbrandt, T., Larsen, T. O., Nordic Natural Product Conference 2017, Odense, DK.

Prediction of Secondary Metabolite Encoding Genes Based on Chemical Structure Analysis. Isbrandt, T., Hoof, J. B., Nielsen, M. L., Gezgin, Y., Nielsen, K. F., Larsen, T. O., Directing Biosynthesis V, 2017, Warwick, UK.

Biosynthesis of acurin A and B in *Aspergillus aculeatus*. Hoof, J. B., Nielsen, M. L., Wolff, P. P., Petersen, L. M., Andersen, L. N., Isbrandt, T., Holm, D. K., Mortensen, U. H., Nødvig, C. S., Larsen, T. O., 29th Fungal Genetics Conference, 2017, Pacific Grove, USA.

Prediction of secondary metabolite encoding genes based on chemical structure analysis. Isbrandt, T., Nielsen, M. L., Hoeck, C., Gezgin, Y., Frandsen, R. J. N., Nielsen, K. F., Larsen, T. O., 9th Joint Natural Products Conference, 2016, Copenhagen, DK.

Prediction of secondary metabolite encoding genes based on chemical structure analysis. Isbrandt, T., Nielsen, M. L., Hoeck, C., Frandsen, R. J. N., Larsen, T. O., 13th European Conference on Fungal Genetics, 2016, Paris, FR.

Production of novel synthetic natural products by engineering of fungal PKS-NRPS hybrids. Nielsen, M. L. ; Isbrandt, T., Mortensen, U. H., Andersen, M. R., Hoof, J. B., Larsen, T. O., 13th European Conference on Fungal Genetics, 2016, Paris, FR.

Exploiting fungal cell factories for pigment production. Tolborg, G., Isbrandt, T., Rasmussen, K. B., Thrane, U., Workman, M., DTU's Sustain Conference 2015, 2015, Technical University of Denmark (DTU), Lyngby, DK.

Other

Controlled production of pink, red, orange and yellow natural pigments in Talaromyces atroroseus. Tolborg, G., Isbrandt, T., Larsen, T. O. & Workman, M., Patent, Priority date: May 8th 2017.

Abbreviations

A domain	Adenylation domain
A/Ala	Alanine
ACP	Acyl carrier protein
ADEQUATE	Adequate Sensitivity Double Quantum Spectroscopy
AT	Acyl transferase
AX	Anion exchange
BGC	Biosynthetic Gene Cluster
BLAST	Basic Local Alignment Search Tool
BPC	Base Peak Chromatogram
C domain	Condensation domain
C/Cys	Cysteine
C18	Octadecyl
C6-Ph	Phenyl hexyl
C8	Octyl
CoA	Co-enzyme A
COSY	Correlation Spectroscopy
CRISPR-Cas9	Clustered Regularly Interspaced Short Palindromic Repeats - CRISPR-associated protein 9
CX	Cation exchange
D/Asp	Aspartic acid
Da	Dalton
DAD	Diode Array Detection
DH	Dehydratase
DMAPP	Dimethylallyl Pyrophosphate
E/Glu	Glutamic acid
edHSQC	Multiplicity Edited Heteronuclear Single Quantum Correlation
EIC	Extracted Ion Chromatogram
ER	Enoyl Reductase
ESI	Electrospray Ionisation

EtOAc	Ethyl acetate
F/Phe	Phenyl alanine
FAS	Fatty Acid Synthase
G/Gly	Glycine
H/His	Histidine
H2BC	Heteronuclear 2-Bond Correlation
HMBC	Heteronuclear Multiple Bond Correlation
HPLC	High Performance Liquid Chromatography
HRMS	High Resolution Mass Spectrometry
I/Ile	Isoleucine
IEX	Ion exchange
IPP	Isopentenyl Pyrophosphate
K/Lys	Lysine
KR	Ketoreductase
KS	Ketosynthase
L/Leu	Leucine
LR-HSQMBC	Long-Range Heteronuclear Single Quantum Multiple Bond Correlation
M/Met	Methionine
m/z	Mass-to-charge ratio
MeOH	Methanol
MIC	Minimum Inhibitory Concentration
mRNA	Messenger Ribonucleic Acid
MS	Mass spectrometry
MT	Methyl transferase
N/Asn	Asparagine
NMR	Nuclear Magnetic Resonance
NOESY	Nuclear Overhauser Effect Spectroscopy
NP	Normal phase
NRP	Non-ribosomal Peptide
NR-PKS	Non-reducing Polyketide Synthase

NRPS	Non-ribosomal Peptide Synthetase
P/Pro	Proline
PCP	Peptidyl Carrier Protein
PFP	Pentafluorophenyl
PK	Polyketide
PKS	Polyketide Synthase
PKS-NRPS	Polyketide Synthase-Non-ribosomal Peptide Synthetase
PR-PKS	Partially Reducing Polyketide Synthase
Q/Gln	Glutamine
qTOF	Quadropole time-of-flight
R domain	Reduction domain
R/Arg	Arginine
RP	Reverse phase
S/Ser	Serine
SAM	S-Adenosyl methionine
T/Thr	Threonine
TE	Thioesterase
TIC	Total Ion Current
TOCSY	Total Correlation Spectroscopy
UHPLC	Ultra-High Performance Liquid Chromatography
USER	Uracil-specific excusion reagent
UV-VIS	Ultraviolet-visual
V/Val	Valine
W/Trp	Tryptophan
WT	wildtype
Y/Tyr	Tyrosine

Contents

Introduction

Fungi

The biodiversity of the fungal kingdom is astonishing, and there is currently more than 80000 described species, and estimated to be in excess of 1.5 million species yet to be discovered.[1] The phylum Ascomycota is just a small part of the whole fungal kingdom, and here the genera Aspergilli and Penicillia are found, of which there are more than 180 and 300 different species respectively.[2,3] Thanks to millions of years of specialisation, each species possess a different array of secondary metabolites, used by the fungi to increase their survival chances in nature.[4]

Fungi are ubiquitous eukaryotic organisms consisting of mushrooms, yeasts, and moulds, that take up a very important part of the ecosystem, as they are the main decomposers of organic material. In addition to their importance in the ecosystem, they have for centuries been used by humans for production of fermented foods such as bread, beer, wine, and cheese.[5]

One type of fungi are moulds, or filamentous fungi, with some of the most famous genera being the Penicillia and Aspergilli (Figure 1). Filamentous fungi are often prolific producers of so called secondary metabolites, that are used by the organism as defence against competing species, as communication molecules, as attractants to help distributing spores, or as protection against the environment.[6,7]

Figure 1. Microscope image of an Aspergillus terreus conidiophore.

Humans have through time found some of these secondary metabolites useful, with the antibiotic penicillin, and the cholesterol lowering drug lovastatin, being some of the most famous. In addition to beneficial compounds, many fungi also produce mycotoxins, which can be either hepatotoxic, nephrotoxic, carcinogenic, or otherwise harmful. Aflatoxin B1, ochratoxin A, fumonisin B1, and citrinin are all examples of harmful fungal secondary metabolites.[8]

In recent years, genome sequencing have become an increasingly affordable technology[9], and with the use of bioinformatics, the full metabolic potential of the various fungal species is gradually being uncovered. Secondary metabolites are produced by a range of different biosynthetic enzymes, usually found in so-called biosynthetic gene clusters (BGCs). These are all encoded in the genome of the organism and constructed via translation of messenger RNA (mRNA) into protein (Figure 2). From analysing large sets of data, it appears that many fungi have the potential to produce a range of so far undiscovered secondary metabolite, encoded in silent genes, not expressed under laboratory conditions.[10]

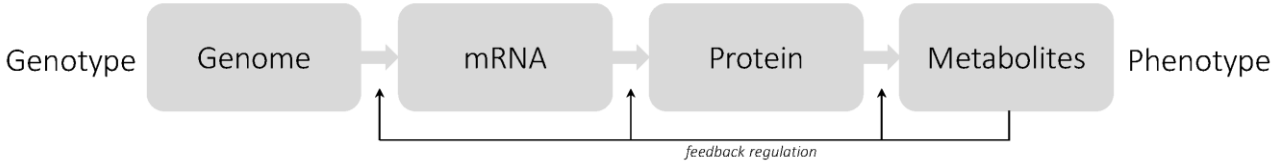


Figure 2. Simplified illustration of the linkage between genes and secondary metabolites. Inspired by Smedsgaard & Nielsen (2003)[11].

Fungal Secondary Metabolites

Microorganisms, such as fungi, are an important source for bioactive molecules, and have for decades been the principal component for discovery of new pharmaceuticals.[12,13] From 1981 to 2014 half of all approved drugs, were either natural products, natural product derivatives, or contained pharmacophores inspired by or mimicking natural products. Figure 3 illustrates the distribution of approved pharmaceuticals based on their origin from 1981 to 2014[13].

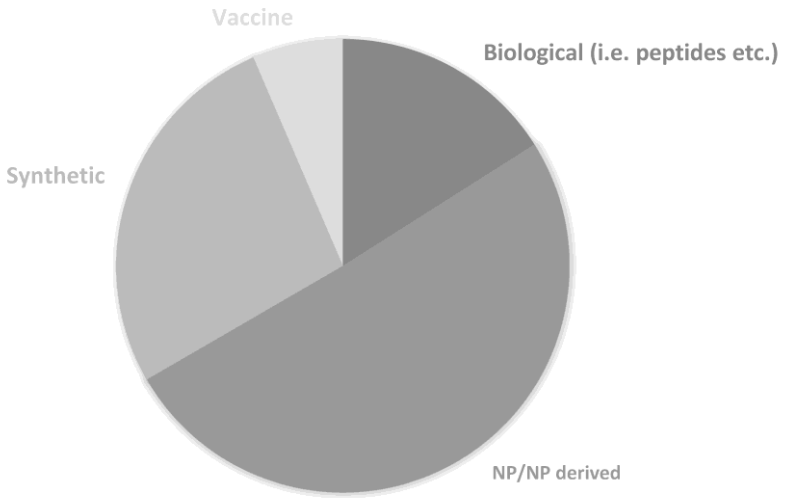


Figure 3. Distribution of approved drugs from 1981 to 2014 by source, based on Newman et al. (2016)[13].

One of the most successful natural products used in the pharmaceutical industry, is the cholesterol lowering drug lovastatin, marketed by Merck under the trade name Mevacor. Lovastatin is produced by a range of

different fungi, here among the filamentous fungus *Aspergillus terreus*, but also macro fungi such as Oyster mushrooms.[14,15]

Aside from pharmaceutically relevant compounds, natural products can also be food ingredients such as flavouring ingredients or colourants. During the last decade or so, consumers have been requesting a higher amount of food additives to be natural, and especially colouring agent for use in food products, have been a concern.[16,17] In Asia, the genus of *Monascus* has for centuries been used to produce red rice koji, or red yeast rice, a type of rice inoculated with the filamentous fungus *Monascus purpureus*, to give a distinct red colour to the food. However, the fungus is also able to produce the mycotoxin citrinin, and its use in Europe and North America is therefore limited.

Biosynthesis of Fungal Secondary Metabolites

The fungal metabolism is generally divided into two main branches: The primary and the secondary metabolism. The primary metabolism is the essential mechanisms for life, such as uptake and digestion of nutrients, cell growth, and reproduction, and are shared by virtually all fungi. The secondary metabolism and metabolites, are in contrast not vital for sustaining life, but a set of tools used to give a species an edge against competitors in the environment, and these are often unique to certain genera or species.

Secondary metabolites can be divided into groups, depending on their biosynthetic origin, i.e. the enzymes used to make them. The major classes are:

- Polyketides,
- Non-ribosomal peptides,
- Terpenes, and
- Shikimic acid derivatives

Compounds from each of the four groups are also constructed from different types of building blocks, and can sometimes be combined into hybrid products resulting in an enormous chemical diversity.

Polyketides

The one of largest and most diverse classes of fungal secondary metabolites are the polyketides (PKs). These compounds are built from ketide units typically originating from acetic acid and malonic acid. Polyketides are sometimes classified based on the number of carbons in the backbone: Eight carbons is a tetraketide, ten carbons is a pentaketide, twelve is a hexaketide, and so forth. In order to understand the biosynthesis it is necessary to consider the enzymatic machinery involved, the polyketide synthases (PKSs).[18]

PKSs can be divided into three groups; type I, type II and type III. Type I PKSs are huge enzymes consisting of either several modules each responsible for one extension of the polyketide chain (modular PKSs), or a single module capable of iteratively extending the polyketide chain (iterative PKSs). Type II PKSs consist of multiple smaller enzymes interacting to form a functional complex working in an iterative fashion. Type III PKSs are made from a dimer of two ketosynthase (KS) domains and do not contain the normal acyl carrier protein (ACP) found in type I/II PKSs.[19–21] Most PKSs found in fungi, are iterative type I and these are the ones that will be described in more detail.

The core of an iterative type I PKS, consists of three domains: a ketosynthase (KS), an acyl transferase (AT), and an acyl carrier protein (ACP) domain. The KS domain is responsible for catalysing the C-C bond formation between the starter unit, typically acetyl-coenzyme A (CoA), or the growing polyketide chain, and the extender units (typically malonyl-CoA)) via Claisen condensations. The AT domain works as a loading unit, by accepting either starter or extender units, and feeding these to the polyketide chain attached to the APC domain. Finally, the ACP domain is where the polyketide chain is tethered, and is responsible for transporting it to the various catalytic domains within the enzyme. The extension of the polyketide happens via a Claisen condensation between the polyketide attached to the ACP and extender unit fed by the AT, and is driven by the loss of CO_2. [19–21]

The three domains responsible for the polyketide extension represents only the most simple of PKSs, and additional domains are often present, able to carry out attach methyl groups, or do reductions of the carbonyl groups, as the chain grows. These are methyltransferases (MT), ketoreductases (KR), dehydratases (DH), and enoyl reductases (ER). Depending on how many of these domains are active in a given iteration, a different number of methyl groups can be attached to the backbone, and the carbonyls can be reduced to either a hydroxyl groups, alkenes, or all the way to alkanes. Depending on the degree of reduction, a PKS is called either non-reducing, partially reducing, or highly reducing. After one iteration, the product can either undergo another round of extension, or be released by a thioesterase (TE) domain.[19–21]

In iterative PKSs, prediction of the number of extensions and to what degree a molecule will be reduced or methylated, is not yet possible, and information about the domains in a PKS will only provide limited information about the final structure. Figure 4 shows a schematic representation of an iterative type I PKS.

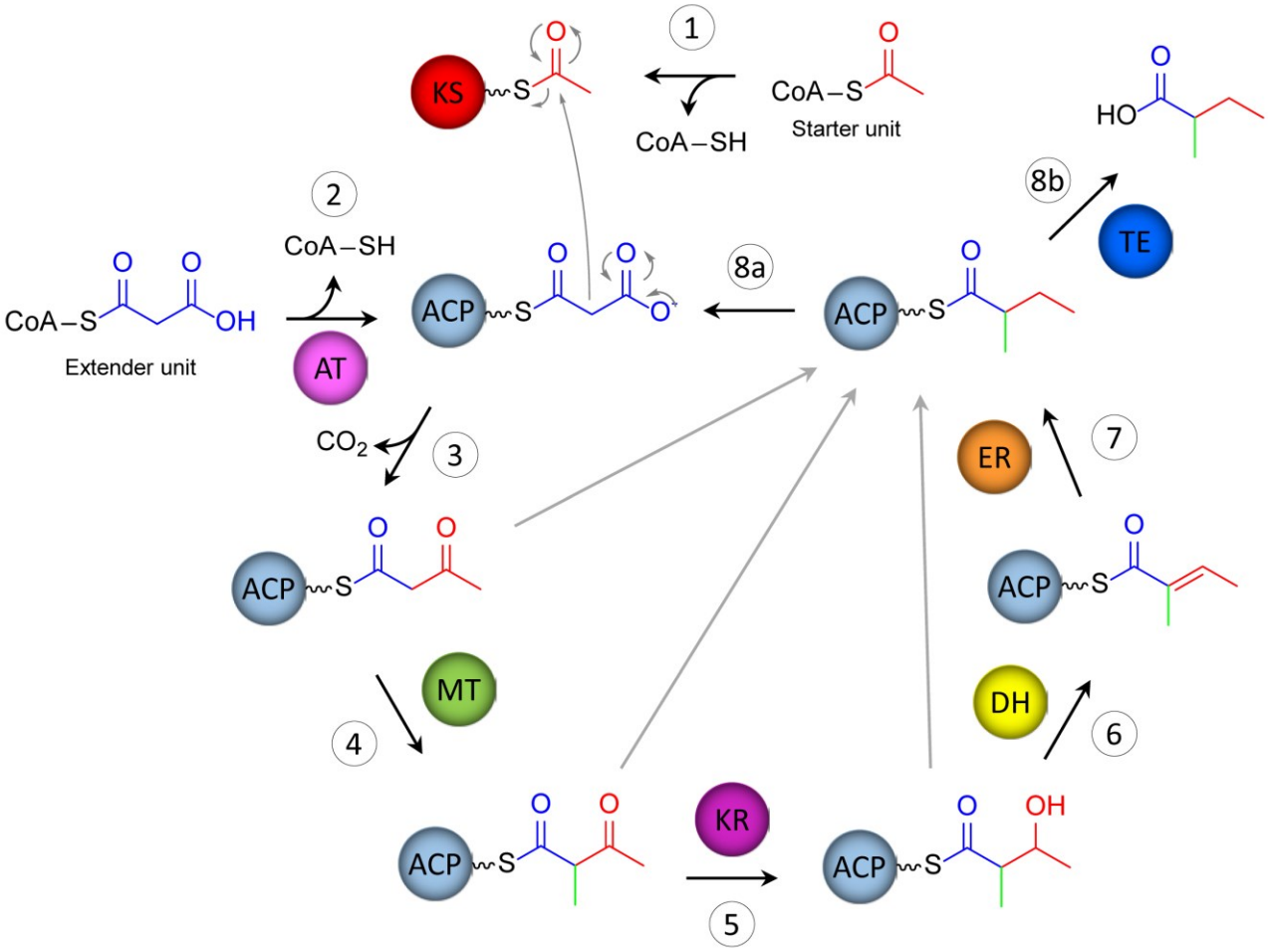


Figure 4. Schematic representation of polyketide biosynthesis. Based on Staunton & Weissman (2001)[19].

In addition to the PKS itself, enzymes encoded by genes in the same biosynthetic gene cluster can further alter the final structure of the product. These are referred to as tailoring enzymes and can have various activities, such as additional reductions, oxidations, cyclysations, or transferases, to mention a few.[22]

An example of a non-reduced polyketide is the pentaketide citrinin (Figure 5). The molecule is made from acetyl-CoA as the starter unit, and four malonyl-CoA as extender units, and has been heavily methylated. Positions three, five and seven are all methylated, and one of the methyl groups has been oxidised all the way to a carboxylic acid. In addition to the methylations, the molecule has cyclised to give a bicyclic heteroaromatic system.[23]

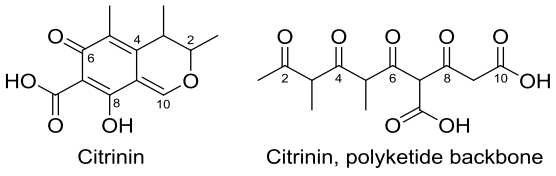


Figure 5. Chemical structure of the mycotoxin citrinin and its non-cyclised polyketide backbone.

Based on the genetic origin of polyketides, it is evident that there is almost unlimited possibilities as to the structures of this class of molecules, depending on the chain length, the reduction and methylation pattern, the cyclisation, and the tailoring.

Non-ribosomal peptides

Non-ribosomal peptides (NRPs) are small peptides made by large modular enzyme complexes called non-ribosomal peptide synthetases (NRPSs) and include compounds such as penicillin G, meleagrin, and malformin C (Figure 6). Being made from amino acid, the potential diversity for these compounds is also remarkable. Each module, consists of several domains, each responsible for a specific task when assembling the product. The domains are adenylation (A) domians, peptidyl carrier protein (PCP) domains, condensation (C) domains, and thioesterase (TE) domains, and additional tailoring domains, such as epimerase, oxidation/reduction, or methyltransferase domains are often present in connection with the NRPS.[18,24,25]

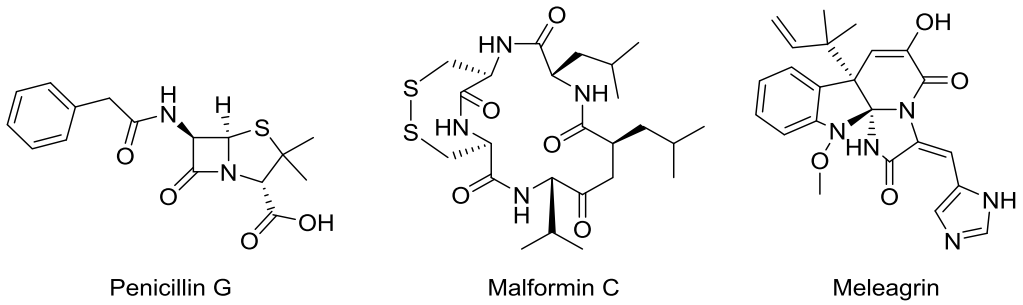


Figure 6. Three different NRPS products, displaying the diversity within the compound class.

The A domain is responsible for selecting the amino acid to be incorporated and resembles the polyketide AT domain, by feeding amino acids to the PCP domain. Unlike AT domains in PKSs, which only recognise a select number of substrates, the NRPS A domains are able to recognise many different substrates, as all natural amino acids, as well as non-proteinogenic amino acids can be incorporated into NRPs. The PCP domain is the NPRS analogue of the ACP, and is where the growing peptide is tethered during biosynthesis. In order to extend the peptide, the C domain catalyses a condensation reaction to form the peptide bonds between the individual amino acids. Since the NPRS is modular, each extension is performed by separate A, PCP, and C domains, and the growing peptide is handed from one PCP domain to the next during the condensation steps. The final step in the biosynthesis is release of the product. This is usually done by a thioesterase (TE) domain, and can be either a simple hydrolysis to give a linear peptide, or a cyclisation resulting in a cyclic peptide. In addition to the domains responsible for constructing the backbone of the molecule, domains able to undertake reactions such as epimerisation, specific cyclisations, methylations, or different release mechanisms are also sometimes found in NRPSs, and tailoring enzymes able to further modify the NRPS product are also common.[18,24] Figure 7 is a schematic representation of the function of an NPRS.

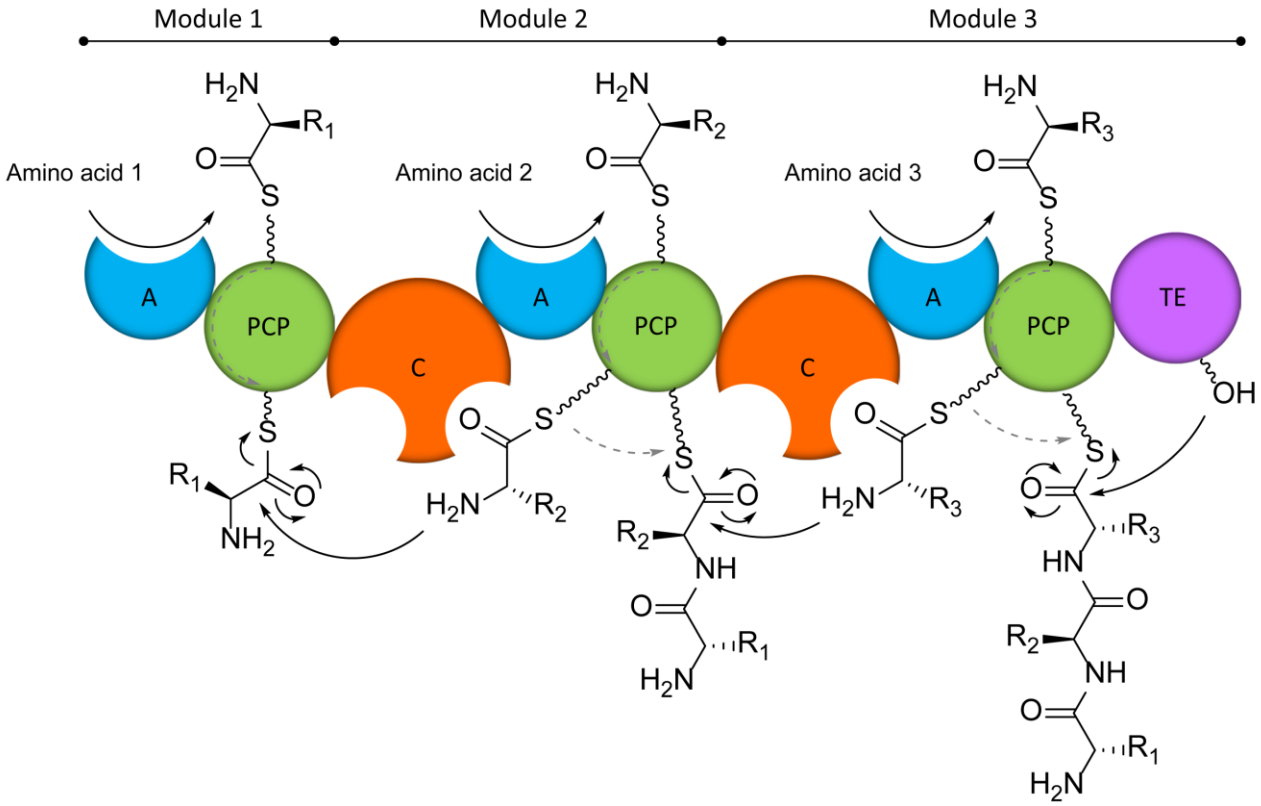


Figure 7. Schematic representation of non-ribosomal peptide biosynthesis. Based on Winn et al. 2016[24].

Terpenes

The third major class of natural products is the terpenes. While these compounds might be less described, compared to polyketides and NRPs, there is still a significant amount of fungal-derived terpenes, such as volatile terpeniods, but most of the biosynthetic studies of these are from within the last decade.[26,27]

Terpenes are built from isoprene units derived from either isopentenyl pyrophosphate (IPP) or dimethylallyl pyrophosphate (DMAPP) (Figure 8). The common feature of these two precursors is the five carbon atoms, and the backbone of terpenes therefore usually contain a number of carbons equal to a multiple of five. Naturally, various modifications can also be done to the molecules after biosynthesis, resulting in either loss or gain of one or more carbon atoms. Depending on the number of isoprene units used for the backbone, terpene are classified as either hemiterpenes (one isoprene unit), monoterpenes (two isoprene units), sesquiterpenes (three isoprene units), diterpenes (four isoprene units), and so on.[26,27]

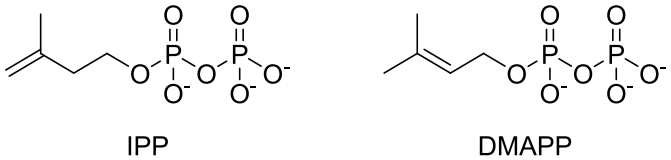


Figure 8. Isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP). The building blocks for terpene biosynthesis.

Some of the most well-known terpenoids, are steroids, such as cholesterol or its fungal equivalent ergosterol, and pigments, such as the plant derived β-carotene. Other fungal derived terpenes include the mycotoxin T-2 toxin produced by species within the genus *Fusarium*.

Shikimic acid derivatives

Fungal compounds originating from the shikimic acid pathway are not as common as PKS, NRPS, or terpenoid derived compounds, but are highly abundant in plants where they function as pigments, structural components or as signalling molecules.[28] Shikimic acid is the precursor for many aromatic amino acids, and compounds such as anthocyanins originate from this compound. The starting unit, shikimic acid, is modified into 4-coumaroyl-CoA, via phenylalanine, cinnamic acid and coumaric acid, and then coupled to several malonyl- and acetyl-CoA units.[28,29] Additional steps lead to the final anthocyanin product, which depending on the exact modifications and surrounding pH can have colours ranging from yellow to red and even blue.

Compounds of mixed biosynthetic origin

In addition to the "pure" polyketides, non-ribosomal peptides and terpenes, many secondary metabolites are derived from combinations of more than one of these three major groups, further expanding the biosynthetic repertoire of each species.

When genes encoding PKSs and NRPSs are found in the same gene cluster, these are usually referred to as PKS-NRPS hybrids, and work by attaching one or more amino acid to the polyketide chain. Some PK-NPR hybrids such as the bacterial immunosuppressant compounds cyclosporine or rapamycin are built from PKs attached to small peptides.[30,31] Examples of fungal PK-NRPs are cytochalasin E[32], myceliothermophin E[33], and talaroconvolutin A[34], shown in Figure 9.

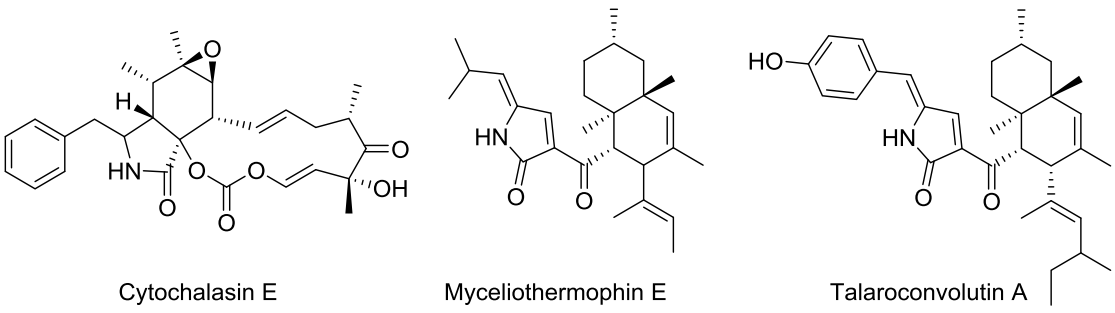


Figure 9. Three structurally related PKS-NRPS products, each originating from different fungal species.

Combinations of PKs or NRPs and terpenes are called meroterpenois and are also common as fungal secondary metabolites. Territrem B, Yanuthone D and meleagrin are examples of meroterpenoids originating from PK and terpene (territrem B and yanuthone D), and NRP and terpene (meleagrin). In addition, territrem B also contain a shikimic acid moiety, which is the precursor for aromatic amino acids in microorganisms and plants.

Bioinformatics and molecular biology in biosynthetic pathway elucidation

The biosynthetic enzymes used the make secondary metabolites are encoded in genes within the genome. In fungi these genes are often found in clusters, that is groups of genes which are regulated and expressed together.[35] With recent technological advances, it has become increasingly more affordable to sequence genomes[9], and using various bioinformatics tools much more in-depth information can be obtained about secondary metabolite gene clusters (Figure 10).[36] This information could be the number of a specific type of gene present in a given organism, the types of tailoring enzymes encoded in a gene cluster, or if certain genes are shared between species. Using bioinformatics, engineering of secondary metabolite genes can also be done more targeted, and using molecular biology and various cloning techniques, such as CRISPR-Cas9, genes can be precisely edited within a single organism or transferred between different ones.[37] Additionally, a large amount of so-called silent genes have been identified in fungi, and molecular biology tools can also be used to activate these otherwise non-expressed genes.[38]

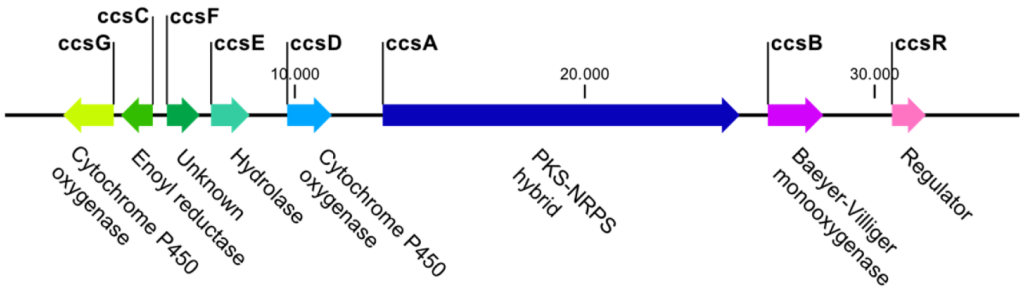


Figure 10. A fungal biosynthetic gene cluster exemplified with the cytochalasin E cluster.

Engineering of secondary metabolite genes can also be used to the elucidate biosynthetic pathway of a secondary metabolite. Once a gene cluster has been identified and confirmed, e.g. by deleting it from the producing organism, knocking out the individual genes and monitoring which secondary metabolites appear or disappear, can be used to determine the order in which the biosynthetic enzymes work together. A few examples of compounds which have had their biosynthetic origin elucidated are cytochalasins[39,40], lovastatin[14,41,42], and *Monascus* pigments[43,44]. These three compounds or compound classes are all polyketides derived, but originate from three different types of PKSs. Cytochalasins are PKS-NRPS hybrid compounds and are made by a highly reducing PKS linked to an NRPS and several tailoring steps; lovastatin is synthesised by a highly reducing PKS with an unusual condensation domain usually found in NRPSs, as well as a small four carbon moiety linked via an ester; and *Monascus* pigments are made by a non-reducing PKS linked to a 3-oxo-fatty acid. In all of the three cases, gene deletion or heterologous expression of either one or more genes was used to suggest the mechanism of formation, and either the genes or the compounds itself have been part of the research described in this thesis.

Discovery of natural products

When analysing extracts from microbial sources, with the goal of discovering novel secondary metabolites, ensuring that resources are not wasted isolating already known compounds is of great importance. One strategy, is dereplication, where various analytical techniques are used in combination to identify compounds based on their chemical properties, such as accurate mass, fragmentation pattern, absorption or emission spectrum, or retention time.[45–48] The most commonly used techniques, are liquid chromatography (LC), UV-VIS spectroscopy, mass spectrometry (MS) and nuclear magnetic resonance spectroscopy (NMR).

Liquid chromatography

Microbial extracts usually contain a plethora of different compounds, and separation of these are essential for efficient application of the various detection methods.

In natural product discovery, liquid chromatography (LC) is the most widely used method for separation of the individual constituents in a sample.[49,50] In a chromatographic column, a mobile and a stationary phase interact with the constituents in the sample, which are retained based on different chemical properties depending on the kind of phases used, e.g. normal phase (NP), reverse phase (RP), or ion exchange (IEX).[49]

In NP chromatography the stationary is polar and is most often made of silica. The mobile phase can be an array of different organic solvents, like chloroform, ethyl acetate, hexane, heptane, toluene, and so on. This causes the more polar analytes to be retained the best, and the non-polar ones to be retained the least. Interactions, such as hydrogen bonding, between the analytes and the stationary phase is responsible for the primary separation in NP chromatography.

Due to challenges with reproducibility, NP chromatography has been outcompeted by RP chromatography, but can still be a useful technique in certain cases, where RP techniques fall short.

In contrast to NP chromatography, RP chromatography uses a non-polar stationary phase and polar mobile phase. The mobile phase is usually a mixture of either methanol or acetonitrile, and water, often buffered with an organic acid such as formic acid or trifluoroacetic acid, or with a base such as ammonium formate, in order to control the pH value. As for NP stationary phases, RP also uses silica as the column material. However the silica in RP columns is derivatised with non-polar groups. Some of the most used RP stationary phases are C18 (octadecyl) and C6-Ph (phenyl-hexyl). These two stationary phase along with biphenyl, C8 (octyl), and PFP (pentafluorophenyl) are shown in Figure 11.[51]

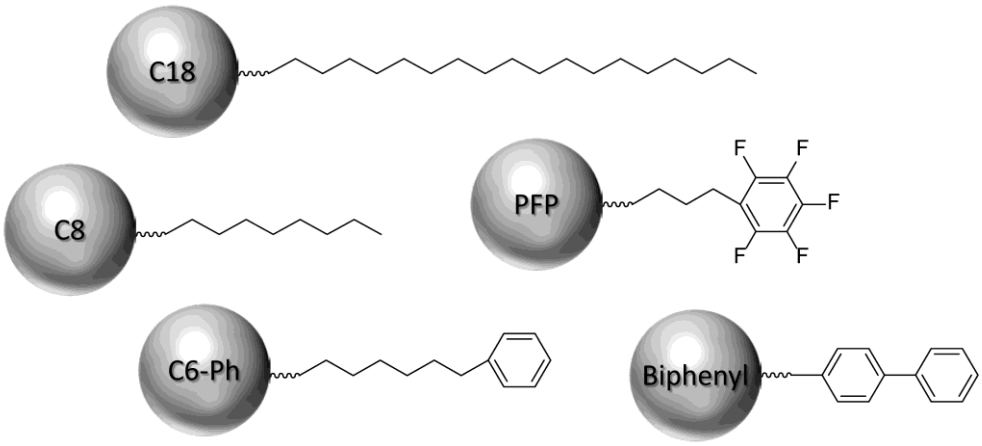


Figure 11. Examples of five different stationary phases used in reverse phase liquid chromatography.

As implied by the name, RP chromatography operates opposite NP, by retaining non-polar analytes more than polar ones. Additionally, selectivity towards different types of compounds can be achieved by using different stationary phases. Aromatic phases like C6-Ph will retain other aromatic compounds more due to π-π-interactions, whereas linear non-polar compounds such as reduced polyketides will be retained better using aliphatic stationary phases, as a result of van der Waals forces. For compounds such as fatty acids, a C18 column might retain the analytes to well, and a C8 could be a better choice, and selectivity towards halogenated compounds can be achieved with a PFP column.

Various parameters will affect the performance of a RP column[49,50], with the particle size of the derivatised silica beads being a main factor. By decreasing the particle size, the surface area with which the analytes can interact is increased, and better separation is accomplished. However, with smaller particles come greater pressure, and high performance LC deals with this, by using high pressure to force the mobile phase though the column. Increasing the length of the column will also provide better separation, but also at the cost of higher pressure. Depending on the task, columns of different sizes are used, and by increasing the diameter, the capacity is increased, allowing more sample to be loaded at a time.

In cases where some analytes have specific properties, such as acidic or basic functionalities, IEX chromatography can provide selectivity towards these compounds, by either retaining acidic compounds (anion exchange, AX) or basic compounds (cation exchange, CX). Depending on the pK_a value of the analytes, either weak or strong IEX materials can be used, and IEX materials with RP properties will give separation based on both RP and IEX properties.

UV-VIS detection

Compounds containing systems of conjugated double bonds are able to absorb ultraviolet and visible light. Depending on the size of the conjugated system, the absorption will shift from the ultraviolet towards the visible region of the electromagnetic spectrum. Additionally, different heteroatoms like oxygen and nitrogen will also alter the way a molecule absorbs light.[52] In Figure 12, three compounds and their absorption spectra are shown, illustrating how different conjugated systems give rise to different spectra.

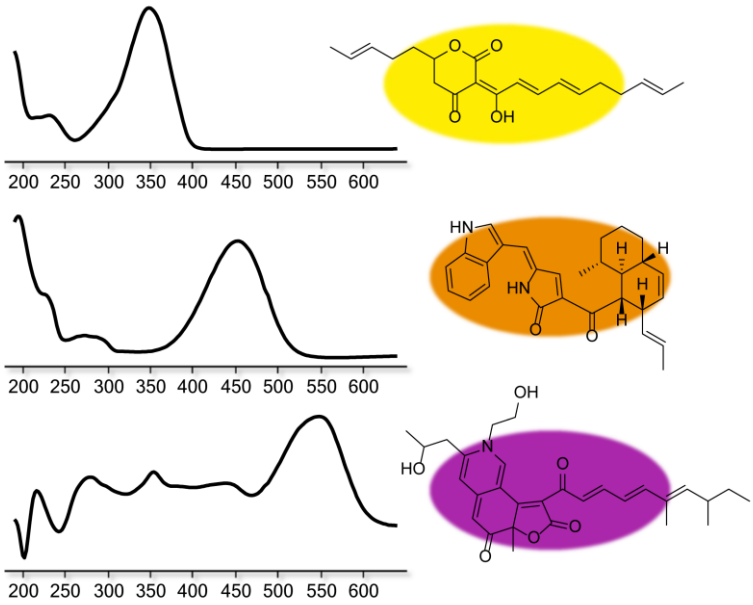


Figure 12. Absorption spectra of three compounds characterised during this project, illustrating the relationship between the size of the conjugated systems and the absorption and observed colour of the compounds.

By exposing analytes to light consisting of wavelengths from ~200-700 nm the absorption spectrum will provide information about the amount of conjugation in the compound. Inversely, having information about the absorption properties of an analyte can help in identifying unknown peaks in a chromatogram. For example, a red pigment will absorb in the area around 500 nm, and monitoring of this area can be used to identify related compounds. Similarly, compounds within certain classes will exhibit related absorption spectra, which can be of help if screening for analogues.

High resolution mass spectrometry and tandem MS

In order to identify an unknown compound, information about the molecular formula is valuable. For unknowns, the number and ratio of carbon, hydrogen, oxygen, and nitrogen atoms, can provide hints towards which biosynthetic family a compound belongs to, and thereby help the structural elucidation. High resolution mass spectrometry (HRMS), is a technique used to measure the accurate mass of compounds, which can then be used to deduce the molecular formula.[53,54]

The formula of an analyte can be calculated using the mass defect. This is a property observed for all elements, and is a result of the different amounts of protons and neutrons in an atom. The carbon isotope ^{12}C has been defined to weigh exactly 12.0000 Da and thus have zero mass defect. Hydrogen and nitrogen have a slight positive mass defect, and oxygen has a slight negative one. In cases where elements with more than one common isotope appear, HRMS can provide information about which and how many of these atoms are in a molecule.[54] In natural product chemistry, the most commonly encountered such elements are chlorine and bromine, and are referred to as A+2 elements because of the mass difference of two Da between the two most abundant isotopes. Both chlorine and bromine exist as two relatively abundant isotopes (^{35}Cl and ^{37}Cl, and ^{79}Br and ^{81}Br), and as a result the isotope patterns for chlorinated and brominated compounds are very distinct and can be used determine which and how many of these atoms are in a molecule.

In MS, the mass of a given molecular ion is measured as the mass-to-charge ratio (m/z). This means that before detection can take place, the molecule needs to be charged. One way of generating ions, is with an electrospray ionisation (ESI) source. ESI is referred to as a soft ionisation technique, meaning that only very little fragmentation takes place in the ion source itself, and that any additional fragmentation has to be done thereafter. The charge is obtained by applying a potential difference between the source and the mass spectrometer, which based on the polarity, can produce either positively or negatively charged ions. In positive mode, the ions are usually proton adducts of the analytes, but other cations such as sodium, potassium, ammonia, or combinations of these are also often observed and are referred to as e.g. $[M+H]^+$ or $[M+Na]^+$. In addition, multiply charged ions, and dimers or trimers can also be formed. Examples of multiply charged ions could be $[M+2H]^{2+}$ or $[M+H+Na]^{2+}$, and examples of dimers and trimers being $[2M+H]^+$, $[2M+Na]^{2+}$, or $[3M+H]^+$. In negative mode, loss of a proton is by far the most commonly observed ion, but formate or chloride adducts sometimes occur. Common in-source fragmentations observed for both positive and negative ESI, is the loss of water or CO_2. In order to assign the correct m/z to a compounds, the mass difference between different adducts are used.[53]

In order to accurately measure the m/z of an ion, several types of MS instruments are available, with quadrupole time-of-flight (qTOF) MS being the one used throughout this project. A qTOF is composed of a quadrupole ion guide which work as an ion selector, a collision cell used to induce fragmentation of the selected ions, and a flight tube used to determine the accurate mass. After ionisation, the ions are directed through a series of ion filters and funnels to the quadrupole. The quadrupole can either let all ions through (full scan) or select specific ions to be fragmented before advancing through the instrument (tandem MS). If an ion is selected for fragmentation, additional energy is applied as it enters the collision cell, and depending on the amount, a range of different fragments is produced by collision with gas molecules in the cell, typically

helium or nitrogen. After the collision cell, the ions, either fragmented or not, travels to the flight tube. The flight tube has a pusler and a detector in one end, and a reflectron in the other. As the ions enter the flight tube, the pulser accelerates the ions perpendicular to their trajectory towards the reflectron. The ions are then reflected back towards the detector where their mass is determined based on the time of the travel from the pusler. The velocity, v, of the ions is determined using $E_{kin} = ½m \cdot v^2$, and since the supplied energy is constant, smaller ions gain a higher velocity and will reach the detector first. In addition to increasing the path and thereby the accuracy, the reflectron also increase the resolution, since ions of the same m/z, but with slightly different distance from the pulser will experience a different amount of energy, and the reflectron corrects for this as the ions with more energy will have an equally longer trajectory.[53] Figure 13 shown a schematic representation of a qTOF.

Tandem MS can be used when searching databases by comparing characteristic fragmentation patterns in order to identify compounds with identical elemental compositions.[47] Furthermore, tandem MS can be used to gain additional structural information about a molecule and to strengthen the structural elucidation of a compound by examining of the fragments and matching these with the structure information obtained from for examples NMR analyses.[55]

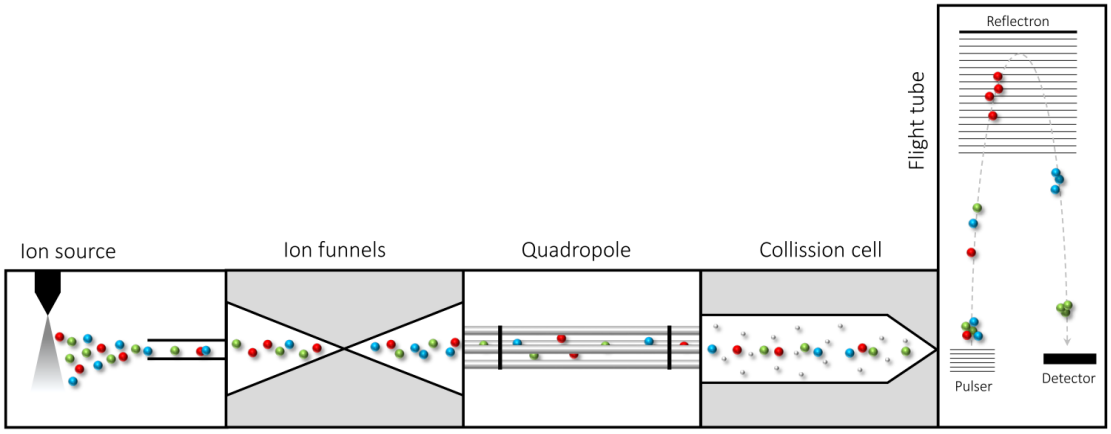


Figure 13. Schematic representation of a quadrupole time-of-flight mass spectrometer.

Nuclear magnetic resonance spectroscopy

Structural elucidation of natural products is not a trivial task, and natural product chemists have an array of tools that can assist when a novel compound is discovered. The most widely used technique is nuclear magnetic resonance (NMR) spectroscopy.[56]

By subjecting a sample to a powerful magnetic field, the atoms in the sample will align their spin with the external field. For spin-½ nuclei such as hydrogen-1 (^{1}H) and carbon-13 (^{13}C), there will be an energy difference depending on the alignment of the spin relative to the external magnetic field and its strength. Depending on

the chemical environment, the individual atoms will be affected differently by the magnetic field, and by exposing them to a radio frequency pulse, each atom will give rise to a different chemical shifts in the NMR spectrum, due to the different frequencies needed to affect each atom. A proton in a CH_3-group will be more shielded by the magnetic field than a lone proton in a CH-group, and even more so than one in a conjugated or aromatic system. This results in the CH_3-proton having the lowest chemical shift, and the aromatic proton the highest. Figure 14 depicts the ranges for common chemical shifts in natural products.[52]

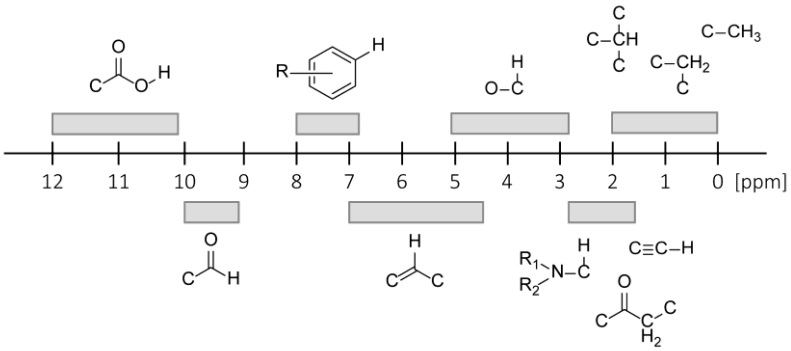


Figure 14. Common chemical shift range for protons in natural products and their type.

These rules also apply to carbon atoms, once again with CH_3-carbons having the lowest chemical shift, CH-carbons having slightly higher, and aromatic and carbonyl carbons giving rise to the highest chemical shifts, as illustrated in Figure 15. In carbon-NMR, the intensity is in most cases much lower than for proton-NMR. This is in part due to the quite low abundance (1.1%) of naturally occurring carbon-13, which is the only stable isotope with useful spin properties (i.e spin ≠ 0). Because of the low abundance of ^{13}C, any experiments involving measurement of carbon requires significantly longer time, and will generate more noise compared to ^{1}H-NMR.

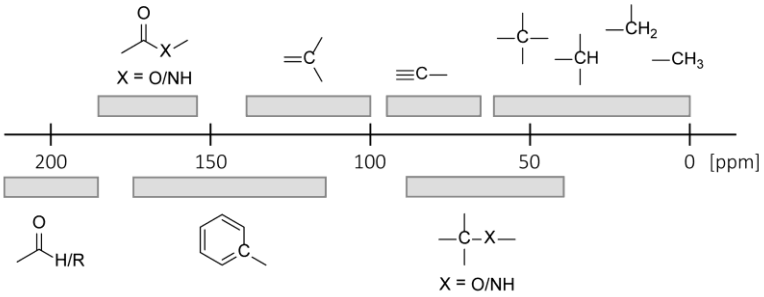


Figure 15. Common chemical shift range for carbons in natural products and their type.

Each signal in an NMR spectrum will, in addition to a frequency, also have an integral corresponding to the amount of protons, as well as a multiplicity. The multiplicity will cause a signal in a spectrum will split up

depending on the number of neighbouring protons, according to the *"n+1 rule"*; i.e. the multiplicity of a peak is equal to the number of chemically equivalent neighbouring protons plus one. If a proton is situated next to a CH_3-group, this will result in a quartet, a proton next to two magnetically different protons will result in a doublet of doublets, and a proton next to a CH_2 and a CH_3 group will result in a sextet. Figure 16 shows examples of these three different types of multiplets. The distance between the peaks in a multiplet is referred to as the coupling constant (*J*), and can for instance be used to determine if protons in an alkene are placed in a *cis* or a *trans* fashion, or if protons in a cyclohexane ring are placed axial or equatorial relative to each other.[56]

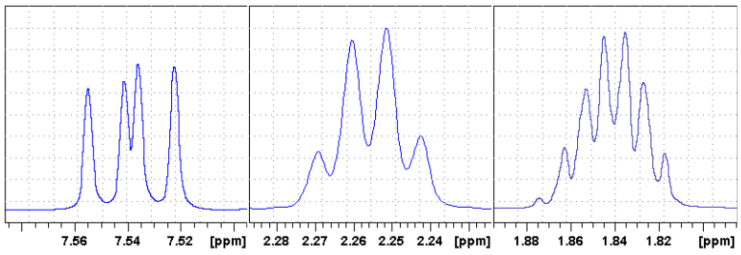


Figure 16. Depiction of three different types of multiplets. A dublet of dublets, a quartet, and a sextet.

Since protons are measured, and samples usually are in the milligram-scale, deuterated solvent are used to dissolve the measured compounds. As deuterium (^2H) has a spin of one, this nuclei in inactive during standard NMR measurements. Depending on the choice of solvent the chemical shift of the signals in a spectrum might change. Furthermore, protic solvents like methanol (CD_3OD) will exchange labile protons in alcohols, amines and carboxylic acids, causing these to not appear in the recorded spectra. This can be prevented by using a non-protic solvent such as acetonitrile (CD_3CN). Other commonly used solvents are dimethylsulfoxide (DMSO-d_6), chloroform (CD_3Cl) or acetone ($(CD_3)_2CO$).

When working with natural products, the molecules are often too complex to be solved using only one-dimensional (1D) ^1H and ^{13}C NMR, and 2D experiments are usually required in order to elucidate the structure of an unknown compound. 2D experiments can be divided into homonuclear or heteronuclear, according to which elements are measured.

Some of the most used homonuclear experiments include *COrrelation SpectroscopY* (COSY)[57,58] and *Nuclear Overhauser Effect SpectroscopY* (NOESY)[59]. In a COSY experiment, information about spin systems, i.e. a series of protons placed on adjacent carbons, can be deciphered. Each signal in a 1D proton experiments will result in a diagonal peak, and protons coupling to each other will share a cross peak, as illustrated in Figure 17. A variation of the COSY experiment is *Total Correlation SpectroscopY* (TOCSY), where all protons in a spin system will appear as cross peaks.

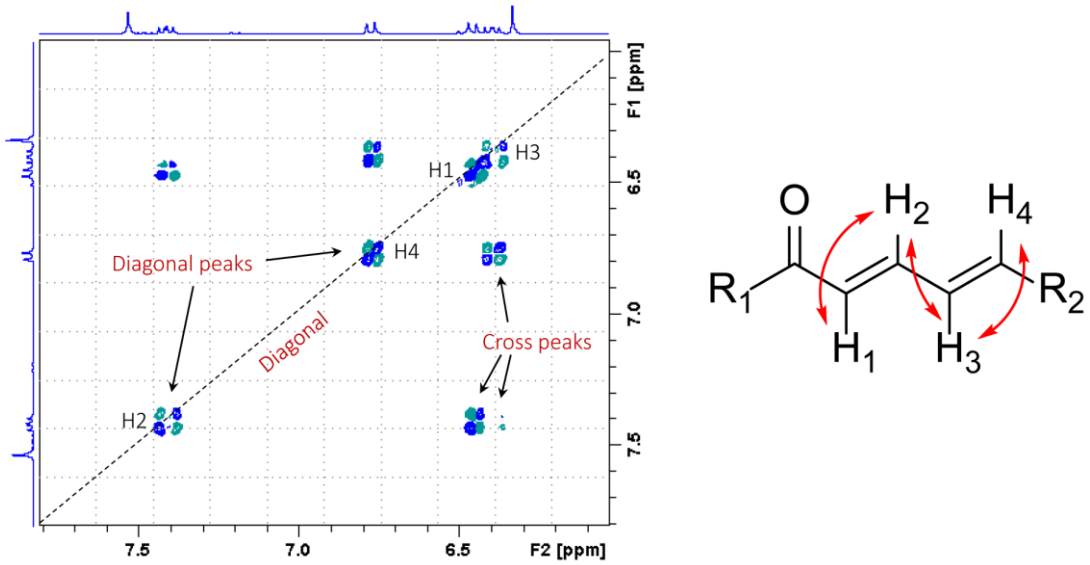


Figure 17. Correlations observed in a COSY experiment.

Where a COSY shows couplings through bonds, a NOESY will show couplings through space. This can be used to determine the relative stereochemistry in molecules with multiple stereocentres. In general, only protons closer than 5 Å apart will give rise to a signal in a NOESY spectrum.

In heteronuclear experiments, the relationship between carbon and proton atoms is measured. There are several different types of experiments, each giving different types of information. A *Heteronuclear Single Quantum Correlation* (HSQC)[60] experiment measures protons and the carbon to which they are directly attached. This means that only protonated carbons will appear in this kind of experiment. However, a variety of the HSQC, is the multiplicity edited HSQC (edHSQC)[61], in which cross peaks in the spectrum will appear as either positive or negative depending on the number of protons attached to a given carbon; CH_2-groups will have opposite sign relative to CH and CH_3-groups. In Figure 18, a single CH_3-group is present along with four CH_2-groups, of which one is diastereotropic, meaning that the two protons experience different electronic environments and therefore have different chemical shifts.

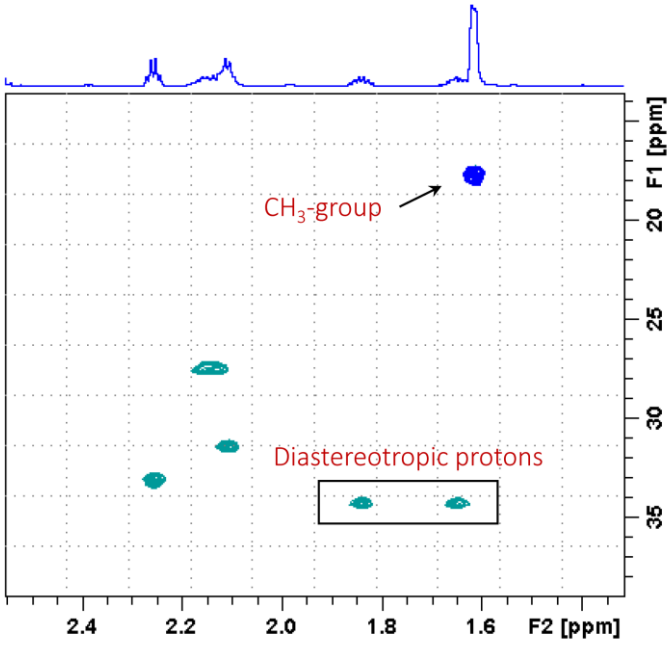


Figure 18. Correlations observed in an edHSQC experiment. The blue peak corresponds to a CH_3-group, and the green ones corresponds to CH_2-groups. The two peaks marked as diastereotropic are two protons attached to the same carbon.

In order to get information about quaternary carbons in a molecule, a *Heteronuclear Multiple Bond Correlation* (HMBC)[62] experiment can be used. In this experiment correlations from a proton to carbons three bonds away is determined, and in some situations, two, four, and sometimes five bond correlations can also be observed in HMBC spectra. Due to the longer correlations observed in a HMBC, it is also an important experiment for linking spin systems together. In Figure 19, a section of a HMBC spectrum is shown along with the correlations in the associated molecule. HMBC experiments can furthermore be optimised towards the longer range couplings (4-5 bond), however with a loss in sensitivity, and specific experiments have been developed to more efficiently and specifically detect four-, five-, and even six-bond correlation (e.g. the *Long-Range Heteronuclear Single Quantum Multiple Bond Correlation* (LR-HSQMBC) experiment[63]).

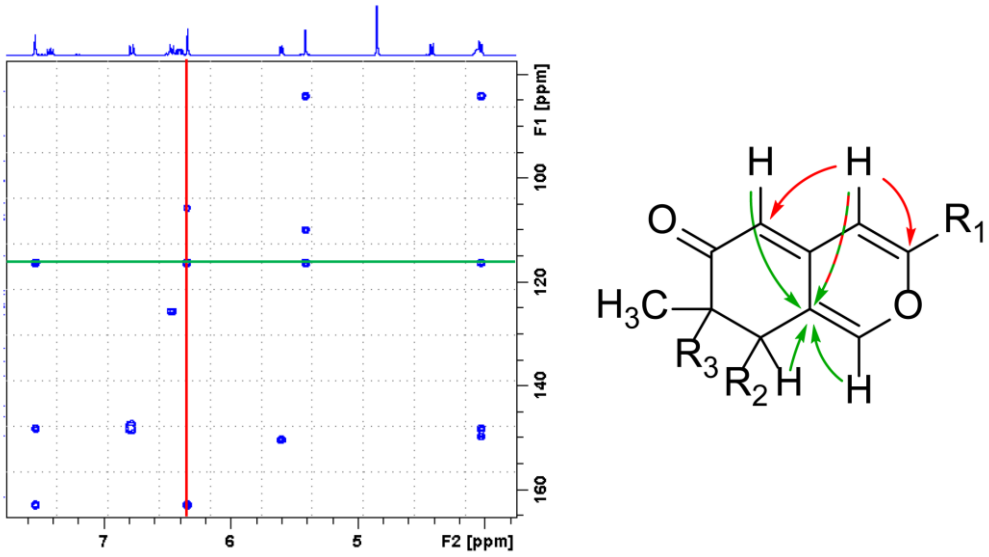


Figure 19. Correlations observed in a HMBC experiment. The green line and arrows shows correlations from several protons to one carbon, and the red line and arrows are correlations from a single proton to several carbons.

In cases with overlap in the proton dimension, a COSY might be insufficient to correctly assign a certain part of a molecule. In these cases, a *Heteronulear 2-Bond Correlation* (H2BC)[64] experiment can provide useful. In a H2BC spectrum, correlations between protons and protonated carbons two bonds away are shown, thereby making the experiment a very good supplement to the COSY, especially in cases where overlap of protons might interfere. Figure 20 shows part of a H2BC spectrum, and the belonging structure.

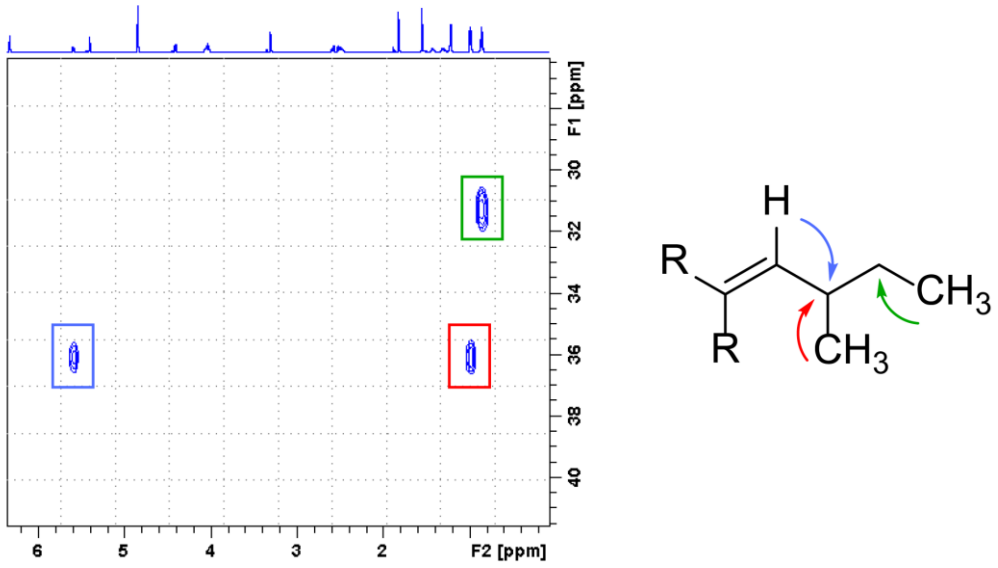


Figure 20. Correlations observed in a H2BC experiment. Each colour represents correlations from different protons to carbons two bonds away.

Sometimes a molecule will contain protons situated in places where the number of correlations will make the HMBC too uncertain, as this experiment does not distinguish between two-, three-, and four bond correlations. Additionally, if the chemical shift of neighbouring carbons are too close, giving a confident assignment might not be possible. In these cases, more specific experiments can be employed, such as ADEQUATE experiments (*Adequate sensitivity Double QUAntum spEctroscopy*)[65]. The 1,1-ADEQUATE experiment can be especially useful, since it specifically shows the correlations two bonds away, via a ^{1}J(H-C) and a ^{1}J(C-C) coupling, between non-protonated, as well as protonated carbons. This provides specific information about which carbons are located next to each other, and can be of great help in cases where other experiments are lacking. However, one major disadvantage of ADEQUATE experiments is very low sensitivity, since a C-C coupling is used.

References

1. Hawksworth, D. L. The magnitude of fungal diversity: the 1.5 million species estimate revisited. *Mycol. Res.* **105,** 1422–1432 (2001).

2. Visagie, C. M. *et al.* Identification and nomenclature of the genus Penicillium. *Stud. Mycol.* **78,** 343–371 (2014).

3. Mousavi, B., Hedayati, M. T., Hedayati, N., Ilkit, M. & Syedmousavi, S. Aspergillus species in indoor environments and their possible occupational and public health hazards. *Curr. Med. Mycol.* **2,** 36–42 (2016).

4. Blackwell, M. & Spatafora, J. W. in *Biodiversity of Fungi* 7–21 (Elsevier, 2004). doi:10.1016/B978-012509551-8/50004-0

5. Lange, L. The importance of fungi and mycology for addressing major global challenges. *IMA Fungus* **5,** 463–471 (2014).

6. Macheleidt, J. *et al.* Regulation and Role of Fungal Secondary Metabolites. *Annu. Rev. Genet.* **50,** 371–392 (2016).

7. Fox, E. M. & Howlett, B. J. Secondary metabolism: regulation and role in fungal biology. *Curr. Opin. Microbiol.* **11,** 481–487 (2008).

8. Bennett, J. W. & Klich, M. Mycotoxins. *Clin. Microbiol. Rev.* **16,** 497–516 (2003).

9. Wetterstrand, K. DNA Sequencing Costs: Data. *Institute, National Human Genome Research* (2017).

at <https://www.genome.gov/sequencingcostsdata/>

10. Kjærbølling, I. *et al.* Linking secondary metabolites to gene clusters through genome sequencing of six diverse Aspergillus species. *Proc. Natl. Acad. Sci.* **115,** E753–E761 (2018).

11. Smedsgaard, J. & Nielsen, J. Metabolite profiling of fungi and yeast: from phenotype to metabolome by MS and informatics. *J. Exp. Bot.* **56,** 273–286 (2005).

12. Newman, D. J. & Cragg, G. M. Natural Products As Sources of New Drugs over the 30 Years from 1981 to 2010. *J. Nat. Prod.* **75,** 311–335 (2012).

13. Newman, D. J. & Cragg, G. M. Natural Products as Sources of New Drugs from 1981 to 2014. *J. Nat. Prod.* **79,** 629–661 (2016).

14. Hendrickson, L. *et al.* Lovastatin biosynthesis in Aspergillus terreus: Characterization of blocked mutants, enzyme activities and a multifunctional polyketide synthase gene. *Chem. Biol.* **6,** 429–439 (1999).

15. Gunde-Cimerman, N. & Cimerman, A. Pleurotus Fruiting Bodies Contain the Inhibitor of 3-Hydroxy-3-Methylglutaryl-Coenzyme A Reductase—Lovastatin. *Exp. Mycol.* **19,** 1–6 (1995).

16. Scotter, M. J. Emerging and persistent issues with artificial food colours: natural colour additives as alternatives to synthetic colours in food and drink. *Qual. Assur. Saf. Crop. Foods* **3,** 28–39 (2011).

17. Amchova, P., Kotolova, H. & Ruda-Kucerova, J. Health safety issues of synthetic food colorants. *Regul. Toxicol. Pharmacol.* **73,** 914–922 (2015).

18. Dewick, P. M. *Medicinal Natural Products*. (John Wiley & Sons, Ltd, 2009). doi:10.1002/9780470742761

19. Staunton, J. & Weissman, K. J. Polyketide biosynthesis: a millennium review. *Nat. Prod. Rep.* **18,** 380–416 (2001).

20. Chooi, Y.-H. H. & Tang, Y. Navigating the fungal polyketide chemical space: from genes to molecules. *J. Org. Chem.* **77,** 9933–53 (2012).

21. Hertweck, C. The Biosynthetic Logic of Polyketide Diversity. *Angew. Chemie Int. Ed.* **48,** 4688–4716 (2009).

22. Weissman, K. J. The structural biology of biosynthetic megaenzymes. *Nat. Chem. Biol.* **11,** 660–670 (2015).

23. Woo, P. C. Y. *et al.* The biosynthetic pathway for a thousand-year-old natural food colorant and citrinin in Penicillium marneffei. *Sci. Rep.* **4,** 6728 (2014).

24. Winn, M., Fyans, J. K., Zhuo, Y. & Micklefield, J. Recent advances in engineering nonribosomal peptide assembly lines. *Nat. Prod. Rep.* **33,** 317–347 (2016).

25. Hur, G. H., Vickery, C. R. & Burkart, M. D. Explorations of catalytic domains in non-ribosomal peptide synthetase enzymology. *Nat. Prod. Rep.* **29,** 1074 (2012).

26. Schmidt-Dannert, C. in *Advances in Biochemical Engineering/Biotechnology* **148,** 19–61 (2014).

27. Matsuda, Y. & Abe, I. Biosynthesis of fungal meroterpenoids. *Nat. Prod. Rep.* **33,** 26–53 (2016).

28. Falcone Ferreyra, M. L., Rius, S. P. & Casati, P. Flavonoids: biosynthesis, biological functions, and biotechnological applications. *Front. Plant Sci.* **3,** 1–15 (2012).

29. Tzin, V. & Galili, G. The Biosynthetic Pathways for Shikimate and Aromatic Amino Acids in Arabidopsis thaliana. *Arabidopsis Book* **8,** e0132 (2010).

30. Fisch, K. M. Biosynthesis of natural products by microbial iterative hybrid PKS–NRPS. *RSC Adv.* **3,** 18228 (2013).

31. Boettger, D. & Hertweck, C. Molecular Diversity Sculpted by Fungal PKS-NRPS Hybrids. *ChemBioChem* **14,** 28–42 (2013).

32. Udagawa, T. *et al.* Cytochalasin E , an Epoxide Containing Aspergillus-Derived Fungal Metabolite , Inhibits Angiogenesis and Tumor Growth. *J. Pharmacol. Exp. Ther.* **294,** 421–427 (2000).

33. Yang, Y. L. *et al.* Cytotoxic polyketides containing tetramic acid moieties isolated from the fungus myceliophthora thermophila: Elucidation of the relationship between cytotoxicity and stereoconfiguration. *Chem. - A Eur. J.* **13,** 6985–6991 (2007).

34. Suzuki, S. *et al.* Antifungal Substances Against Pathogenic Fungi, Talaroconvolutins, from Talaromyces convolutus. *J. Nat. Prod.* **63,** 768–772 (2000).

35. Yi, G., Sze, S.-H. & Thon, M. R. Identifying clusters of functionally related genes in genomes. *Bioinformatics* **23,** 1053–1060 (2007).

36. Luscombe, N. M., Greenbaum, D. & Gerstein, M. What is bioinformatics? A proposed definition and overview of the field. *Methods Inf. Med.* **40,** 346–358 (2001).

37. Nødvig, C. S., Nielsen, J. B., Kogle, M. E. & Mortensen, U. H. A CRISPR-Cas9 System for Genetic

Engineering of Filamentous Fungi. *PLoS One* **10,** e0133085 (2015).

38. Brakhage, A. A. & Schroeckh, V. Fungal secondary metabolites – Strategies to activate silent gene clusters. *Fungal Genet. Biol.* **48,** 15–22 (2011).

39. Qiao, K., Chooi, Y. H. & Tang, Y. Identification and engineering of the cytochalasin gene cluster from Aspergillus clavatus NRRL 1. *Metab. Eng.* **13,** 723–732 (2011).

40. Fujii, R., Minami, A., Gomi, K. & Oikawa, H. Biosynthetic assembly of cytochalasin backbone. *Tetrahedron Lett.* **54,** 2999–3002 (2013).

41. Campbell, C. D. & Vederas, J. C. Mini Review: Biosynthesis of lovastatin and related metabolites formed by fungal iterative PKS enzymes. *Biopolymers* **93,** 755–763 (2010).

42. Xu, W. *et al.* LovG: The thioesterase required for dihydromonacolin L release and lovastatin nonaketide synthase turnover in lovastatin biosynthesis. *Angew. Chemie - Int. Ed.* **52,** 6472–6475 (2013).

43. Chen, W. *et al.* Orange, red, yellow: biosynthesis of azaphilone pigments in Monascus fungi. *Chem. Sci.* **8,** 4917–4925 (2017).

44. Shimizu, T., Kinoshita, H., Ishihara, S., Sakai, K. & Nagai, S. Polyketide Synthase Gene Responsible for Citrinin Biosynthesis in Monascus purpureus Polyketide Synthase Gene Responsible for Citrinin Biosynthesis in Monascus purpureus. *Appl. Environ. Microbiol.* **71,** 1–6 (2005).

45. Hubert, J., Nuzillard, J. M. & Renault, J. H. Dereplication strategies in natural product research: How many tools and methodologies behind the same concept? *Phytochem. Rev.* **16,** 55–95 (2017).

46. Nielsen, K. F., Månsson, M., Rank, C., Frisvad, J. C. & Larsen, T. O. Dereplication of Microbial Natural Products by LC-DAD-TOFMS. *J. Nat. Prod.* **74,** 2338–2348 (2011).

47. Kildgaard, S. *et al.* Accurate dereplication of bioactive secondary metabolites from marine-derived fungi by UHPLC-DAD-QTOFMS and a MS/HRMS library. *Mar. Drugs* **12,** 3681–3705 (2014).

48. Klitgaard, A. *et al.* Aggressive dereplication using UHPLC–DAD–QTOF: screening extracts for up to 3000 fungal secondary metabolites. *Anal. Bioanal. Chem.* **406,** 1933–1943 (2014).

49. Agilent Technologies. *The LC Handbook*. (2016). at <https://www.agilent.com/cs/library/primers/Public/LC-Handbook-Complete-2.pdf>

50. Waters Coorporation. *Beginners Guide to Liquid Chromatography*. (2014). at

<http://www.waters.com/waters/en_US/HPLC---High-Performance-Liquid-Chromatography-Explained/nav.htm?locale=en_US&cid=10048919>

51. Phenomenex. Phase information. (2018). at <https://www.phenomenex.com/Kinetex/Selectivities>

52. Gary M. Lampman, Donald L. Pavia, George S. Kriz & James R. Vyvyan. *Spectroscopy*. (Mary Finch, 2010).

53. Xian, F., Hendrickson, C. L. & Marshall, A. G. High Resolution Mass Spectrometry. *Anal. Chem.* **84,** 708–719 (2012).

54. Pleil, J. D. & Isaacs, K. K. High-resolution mass spectrometry: basic principles for using exact mass and mass defect for discovery analysis of organic molecules in blood, breath, urine and environmental media. *J. Breath Res.* **10,** 12001 (2016).

55. Klitgaard, A., Nielsen, J. B., Frandsen, R. J. N., Andersen, M. R. & Nielsen, K. F. Combining Stable Isotope Labeling and Molecular Networking for Biosynthetic Pathway Characterization. *Anal. Chem.* **87,** 6520–6526 (2015).

56. Breton, R. C. & Reynolds, W. F. Using NMR to identify and characterize natural products. *Nat. Prod. Rep.* **30,** 501 (2013).

57. Marion, D. & Wüthrich, K. Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of 1H-1H spin-spin coupling constants in proteins. *Biochem. Biophys. Res. Commun.* **113,** 967–974 (1983).

58. Rance, M. *et al.* Improved spectral resolution in COSY 1H NMR spectra of proteins via double quantum filtering. *Biochem. Biophys. Res. Commun.* **117,** 479–485 (1983).

59. Kumar, A., Ernst, R. R. & Wüthrich, K. A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochem. Biophys. Res. Commun.* **95,** 1–6 (1980).

60. Bodenhausen, G. & Ruben, D. J. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett.* **69,** 185–189 (1980).

61. Willker, W., Leibfritz, D., Kerssebaum, R. & Bermel, W. Gradient selection in inverse heteronuclear correlation spectroscopy. *Magn. Reson. Chem.* **31,** 287–292 (1993).

62. Cicero, D. O., Barbato, G. & Bazzo, R. Sensitivity Enhancement of a Two-Dimensional Experiment for

the Measurement of Heteronuclear Long-Range Coupling Constants, by a New Scheme of Coherence Selection by Gradients. *J. Magn. Reson.* **148,** 209–213 (2001).

63. Williamson, R. T., Buevich, A. V., Martin, G. E. & Parella, T. LR-HSQMBC: A Sensitive NMR Technique To Probe Very Long-Range Heteronuclear Coupling Pathways. *J. Org. Chem.* **79,** 3887–3894 (2014).

64. Nyberg, N. T., Duus, J. Ø. & Sørensen, O. W. Heteronuclear Two-Bond Correlation: Suppressing Heteronuclear Three-Bond or Higher NMR Correlations while Enhancing Two-Bond Correlations Even for Vanishing 2 J CH. *J. Am. Chem. Soc.* **127,** 6154–6155 (2005).

65. Reif, B. *et al.* ADEQUATE, a New Set of Experiments to Determine the Constitution of Small Molecules at Natural Abundance. *J. Magn. Reson. Ser. A* **118,** 282–285 (1996).

Spectroscopy and spectrometry guided discovery of natural products

Summary of:

Appendix 1: Atrorosins: a new subgroup of *Monascus* pigments from *Talaromyces atroroseus*

Appendix 2: Unique processes yielding pure azaphilones in *Talaromyces atroroseus*

Appendix 3: Establishing novel cell factories producing natural pigments in Europe (In: Biopigmentation and biotechnological implementations)

Appendix 4: New azaphilones from *Aspergillus neoglaber*

Appendix 5: Structure and genetic origin of novel class of polyketide biomarkers from *Aspergillus brasiliensis*, brasenols

Identification of a new group of Monascus pigments from *Talaromyces atroroseus* (Appendix 1-3)

The pigment production of the filamentous fungus *Talaromyces atroroseus* was investigated through controlled cultivation in bioreactors. Chemical analysis using Ultra-High Performance Liquid Chromatography coupled to Diode Array Detection and High Resolution Tandem Mass Spectrometry (UHPLC-DAD-HRMS/MS), led to the identification of a range of known yellow, orange and red compounds, as well as the discovery of several unknown compounds with chromophores absorbing light in the red area of the electromagnetic spectrum (~520 nm). The fermentation broth was extracted using acidic ethyl acetate (EtOAc), and the extract was fractionated with flash chromatography, and using semi-preparative HPLC, a red compound with a mass-to-charge ratio of 500.1912 ($C_{26}H_{29}N_2O_9$) was purified. Structural elucidation using nuclear magnetic resonance (NMR) spectroscopy was used to elucidate the structure shown in Figure 21 (full description of the structural elucidation can be found in Appendix 1). The compound was named atrorosin S, with the S denoting the identity of the amino acid.

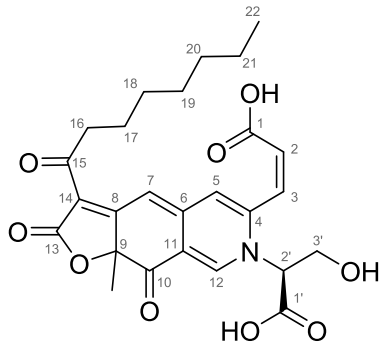


Figure 21. Structure of atrorosin S isolated from T. atroroseus.

Atrorosin S was found to be an azaphilone belonging to the class of *Monascus* pigments, structurally similar to the two known pigments PP-O and PP-V, but differing from these by incorporation of a serine moiety in the isochromene/isoquinoline ring system. Based on the new structure, the extract was reanalysed for similar compounds with incorporation of each of the 20 proteinogenic amino acids. This analysis was able to tentatively identify 14 of the remaining 19 amino acid derivatives in additions to atrorosin S. In order to structurally characterise each of the remaining pigments, shake flask cultivations using each of the amino acids as the sole nitrogen source was carried out. Surprisingly, this resulted in very clean production of single atrorosins, with incorporation of each of the fed amino acid for most of the experiments[1], as shown in Figure 22. From the shake flask cultivations, an additional seven atrorosins were isolated.

[1] No production of the proline analogue was expected or observed, likely due to the lack of a primary amine in the amino acid, needed for incorporation in the azaphilone core.

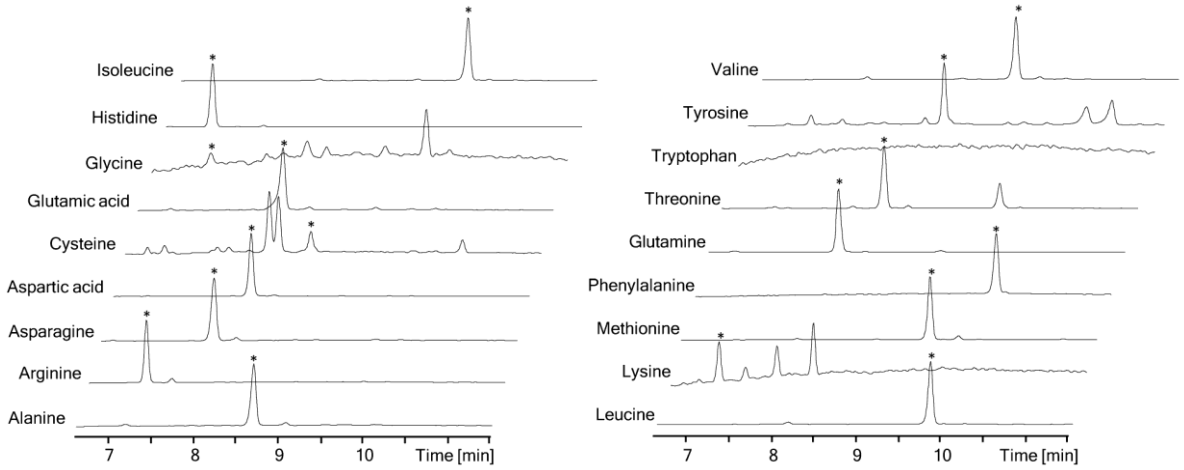


Figure 22. UV-VIS chromatograms (520±10nm) of each of the shake flask cultivations using single amino acids as the sole nitrogen source.

The remaining 11 atrorosins were semi-synthesised from the orange precursor PP-O. For isolation of PP-O, a bioreactor cultivation was performed. This resulted in production of two isomers of PP-O, *cis*- and *trans*-PP-O, differing by the configuration of the double bond between C2 and C3, as shown in Figure 23. By reacting aliquots of *cis*-PP-O with excess of amino acids, it was possible to obtain nature identical versions of each of the remaining atrorosins.

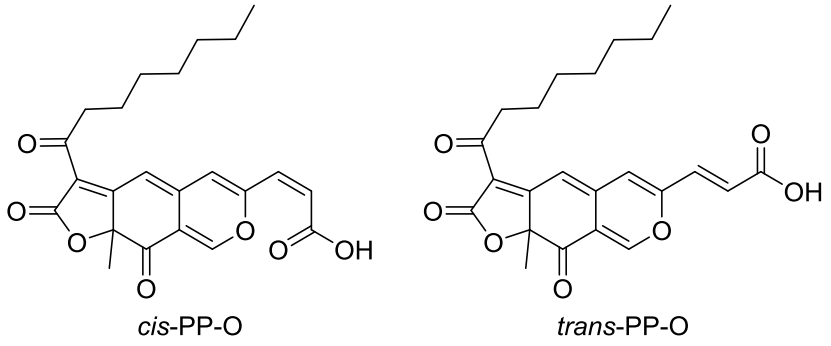


Figure 23. Structure of the two PP-O isomers, cis- and trans-PP-O.

By transferring the knowledge learned from the shake flask cultivations to fermentation in bioreactors, yields of 0.9 g/L of the pigments could to obtained by different cultivation strategies and feeding with single amino acids (Appendix 2).

During the fermentation process, PP-O was observed to be produced as two isomers. However, when fed with amino acids, either as the sole nitrogen source (one-step fermentation), or during the fermentation process (two-step fermentation), the atrorosins were primarily observed as the *cis* isomer. More specifically, the two PP-O isomers were produced in a 3:2 *cis*:*trans* ratio, whereas 99.5% of the atrorosins where very surprisingly produced as the *cis* isomers (Figure 24).

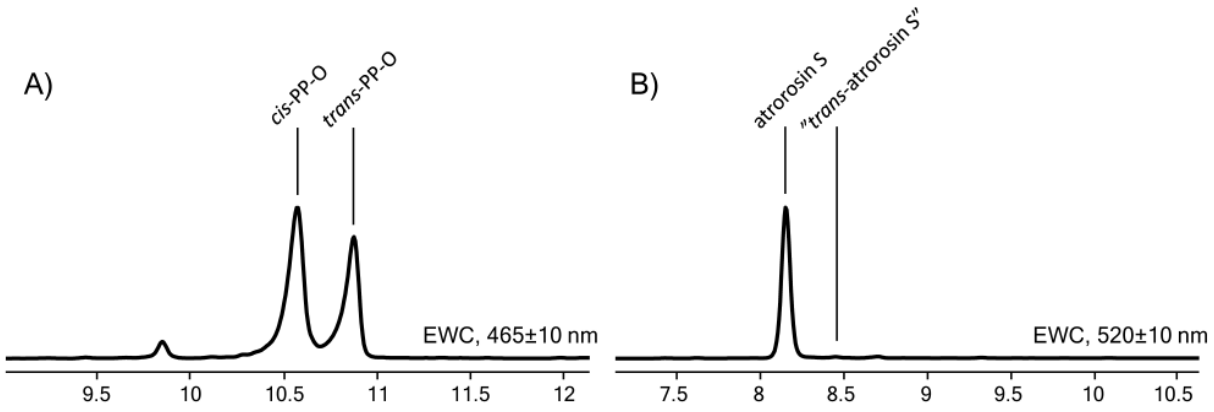


Figure 24. UV-VIS chromatograms showing the change in isomer ratio between PP-O before, and atrorosin S after addition of serine to the bioreactor.

The name azaphilone means "nitrogen loving", and it is well known that compounds within this class is able to react with primary amines.[1] Whether the incorporation of amino acids happens spontaneously or is enzyme catalysed is therefore an intriguing question. Semi-synthesis of atrorosins was possible under acidic (appendix 1) as well as basic conditions (appendix 2), and other studies also report modification of azaphilones such as conversion of the yellow PP-Y into its amino acid derivatives[2]. Several studies have been performed on the effect of nitrogen sources on pigment production in both *Penicillium purpurogenum*[3,4] (=*T. atroroseus*) and *Monascus* species[5,6], but the mechanism by which the incorporation happens remains elusive. Based on the work done during this project, there is still no clear answer, but certain observations seems to suggest that enzymes might play a role. First of all, when cultivated in shake flasks with single amino acids as nitrogen source, not all amino acids performed equally well. From bioreactor studies however, it became apparent that the pH during cultivation plays a huge factor with regard to the amount of pigment produced (appendix 2), and that the non-controlled environment in the shake flasks likely was a factor causing the observed differences. Upon reaction of several amino acids with a mixture of *cis*- and *trans*-PP-O, all experiments resulted in the same ratio of atrorosins and their *trans*-versions, while reaction with a single PP-O isomer only gave the expected atrorosin isomer. When cultivating the fungus in a two-step manner as described in Appendix 2, PP-O was first produced as two isomers, but upon addition of an amino acid, unexpectedly only a single atrorosin was detected. This observation could indicate that the products are either converted into the *cis* form during the transamination reaction, or at some point after the conversion, but with no further investigations the exact reason for the observation of only a single product remains speculative.

Novel azaphilones from *Aspergillus neoglaber* (Appendix 4)

In addition to the atrorosins isolated from *T. atroroseus*, azaphilones from another fungus, *A. neoglaber*, was also investigated in this project (appendix 4). Initially, the fungus was expected to produce a novel arisugacin analogue based on dereplication. However, isolation and NMR studies of the compound quickly showed that

the compound was in fact an azaphilone, related to a group of known compounds called sassafrins[7], and that additional analogues was also produced by the fungus. In total, three compounds were isolated from *A. neoglaber*, two novel sassafrin analogues (sassafrin E and F), and a nitrogen containing derivative named sassafrinamine A. Sassafrin E and F, were composed of the same hexaketide azaphilone core, attached to a second linear hexaketide as the known sassafrins A-C, but differed by an alternate reduction pattern of the linear carbon chain. Sassafrinamine A resembled the sassafrin-like berkchaetorubramine[8], with incorporation of an aminoethanol molecule, but was constructed from the same two twelve-carbon units as the sassafrins, with an additional oxidation forming a double bond between the two conjugated systems in the molecule. This resulted in a violet appearance of the pure compound and a quite unique absorption spectrum. Figure 25 shows the three compounds isolated from *A. neoglaber*, as well as examples of the already known analogues.

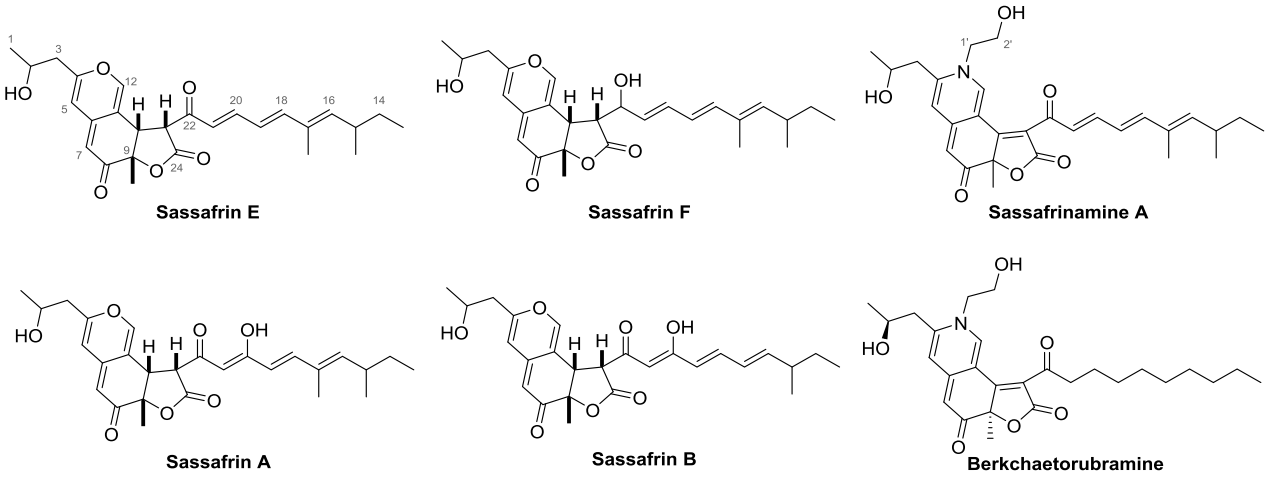


Figure 25. Chemical structures of the three compounds isolated from A. neoglaber in this project (top row), and three known analogues (bottom row).

Structural elucidation of the central part (specifically correct placement of carbons C-6 and C11) of sassafrin E was not trivial, and advanced NMR experiments was necessary for unambiguous assignment of the structure. Using the 1,1-ADEQUATE experiment[9] it was possible to assign the immediate neighbours to the protons, H-5, H-7, H-10, and H-12, in the central part of the molecule, thereby establishing the azaphilone core. Additionally, tandem mass spectrometry and deuterium labelling experiments were used to further confirm the structure. Figure 26 shows the proposed fragmentation pathways of sassafrin E based on the deuterium labelling experiments.

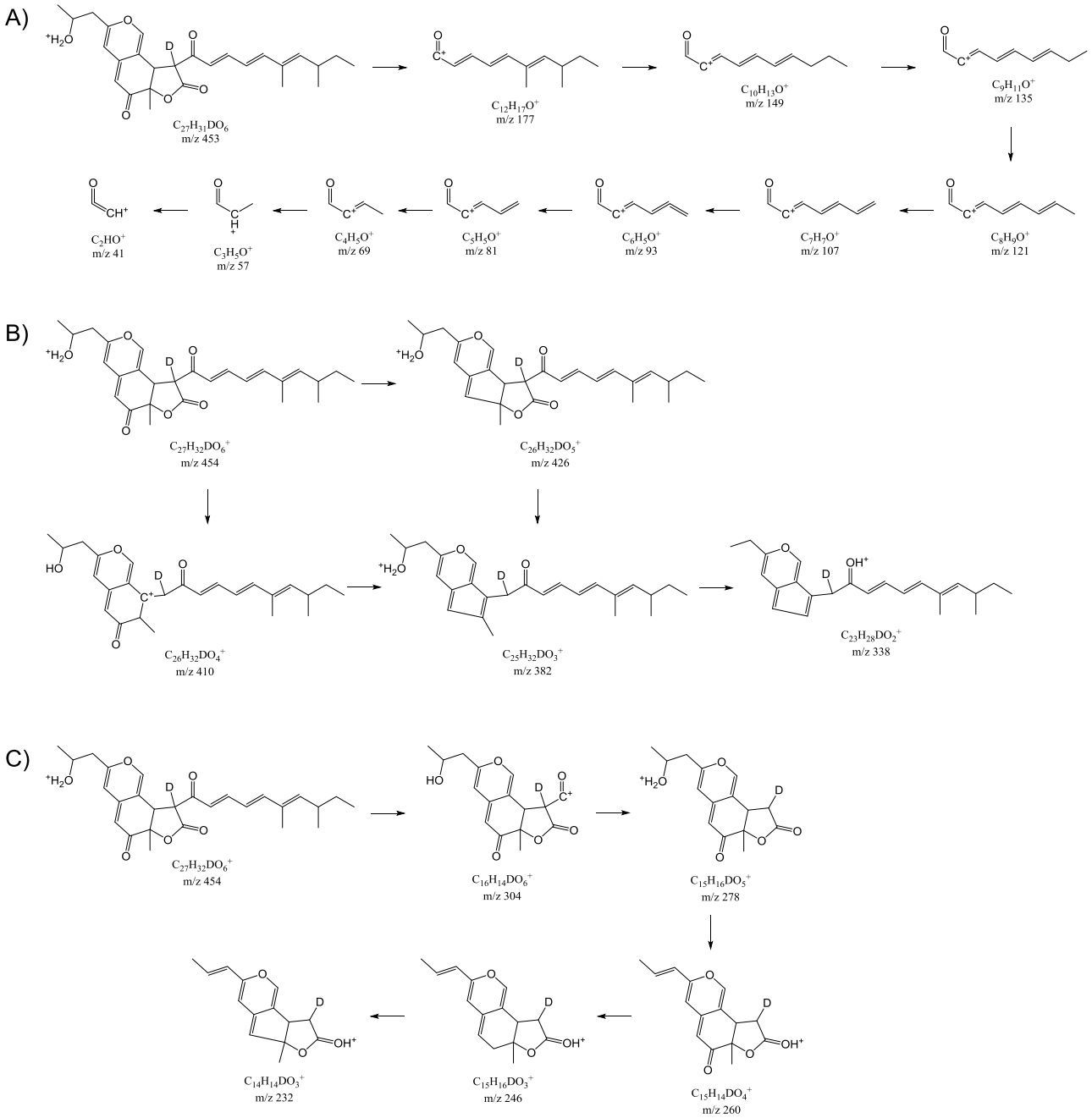


Figure 26. Suggested fragmentation pathways for deuterium labelled sassafrin E.

Spectroscopy guided discovery of a novel class of *Aspergillus* biomarkers (Appendix 5)

The compounds from *T. atroroseus* and *A. neoglaber* were isolated based on dereplication that revealed compounds of interest on the grounds of either recognisable absorption spectra (*Monascus* pigment-like) or molecular formula (a possible arisugacin analogue). Another approach is investigation of unfamiliar features, in search of novel chemistry (Appendix 5). A strain of the black aspergilli *Aspergillus brasiliensis* was found to produce large amounts of a late eluting compound with a unique absorption spectrum (Figure 27B).

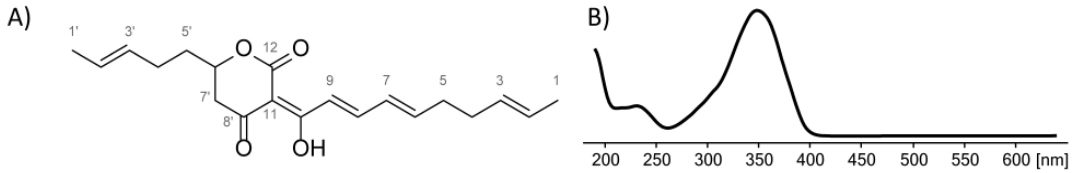


Figure 27. Structure of brasenol A1 (A) and its absorption spectrum (B).

Based on NMR studies, the compound (brasenol A1) was determined to be of polyketide/fatty acid origin, speculated to be made from a twelve-carbon PK unit fused to an eight-carbon 3-hydroxy-fatty acid via a central lactone ring (Figure 27A). Initial bioactivity assays showed mild activity against methicillin-resistant *Staphylococcus aureus* (MRSA) MB5393 (MIC = 28.44 µg/mL), and from searching our in-house database the compound, along with several tentative analogues, was determined to also be produced by two other black Aspergilli, *A. carbonarius* and *A. ellipticus*. Based on these findings, bioinformatics comparison of *A. brasiliensis* and *A. carbonarius* was undertaken with the aim of uncovering the biosynthesis of the compound. By conducting a BLAST search with using ketide synthase (KS) sequence against the protein sequences of *A. brasiliensis*, *A. carbonarius* and the non-producing strain *A. tubingensis*, and comparing the resulting protein sequences, several candidate genes were identified. Figure 28 shows a cladogram of the three strains used for the BLAST analysis, highlighting the candidates with highest similarity between *A. brasiliensis* and *A. carbonarius*.

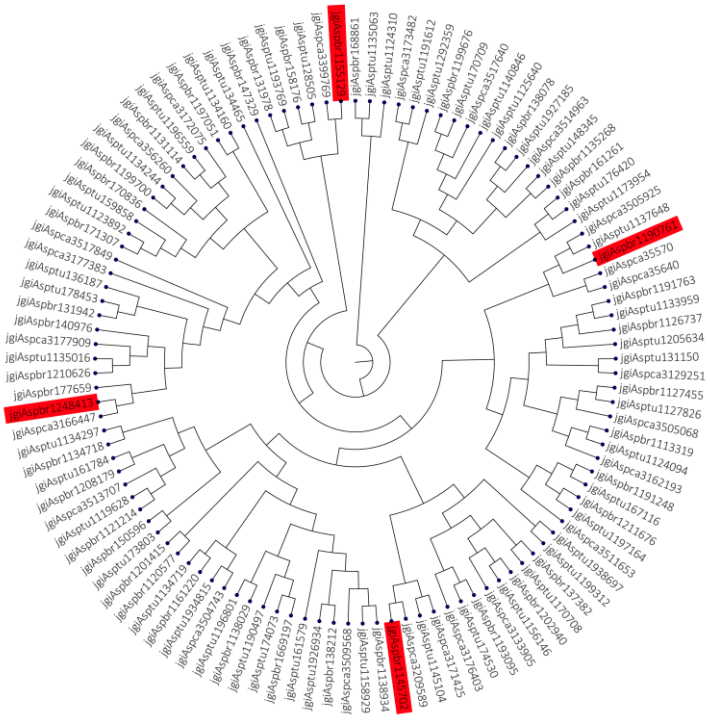


Figure 28. Cladogram of the filtered results from the bioinformatics search for the brasenol PKS gene. The four candidate genes that were selected for deletion have been highlighted.

The candidates identified in *A. brasiliensis* were deleted, resulting in abolishment of brasenol production in only one strain, Abra1_145702Δ (*brsA*). Further analysis of the genes surrounding *brsA* identified putative hydrolase and esterase-like genes, possibly also involved in the biosynthesis. This was confirmed as deletion of the hydrolase, *brsB*, also resulted in cessation of brasenol production and deletion of the esterase, *brsC*, lead to a severe decrease in production. UV chromatograms of the wildtype and the *brsA*Δ, *brsB*Δ and *brsC*Δ strains are shown in Figure 29.

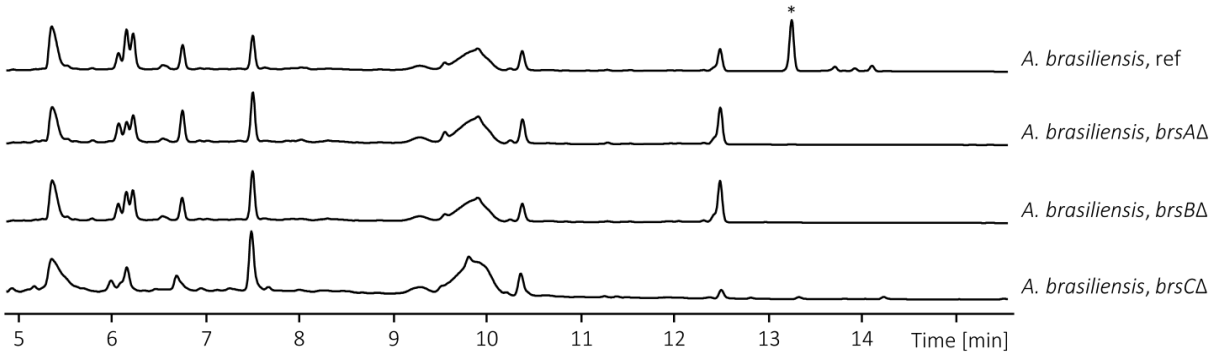


*Figure 29. UV-VIS chromatograms at 350±5nm for the A. brasiliensis wildtype (ref) and the BrsA, BrsB and BrsC deletion strains, showing loss of brasenol production in the brsAΔ and brsBΔ mutants, and significantly lower production in the ΔbrsC strain. The peak corresponding to brasenol A1 has been marked with *.*

Based on the hypothesis that brasenol A1 was made of two individual carbon chains, *brsA* and *brsB* were heterologously expressed in three different fungi, *A. nidulans*, *A. sydowii* and *A. oryzae.* Expression of either of genes by themselves did not show any change in the secondary metabolite profile of the heterologous hosts. However, when *brsA* and *brsB* were expressed together, significant production of brasenol A1, as well as the analogues also seen in *A. brasiliensis*, was observed (Figure 30. Heterologous expression of the two genes brsA and brsB in the three hosts, A. nidulans, A. sydowii, and A. oryzae.Figure 30).

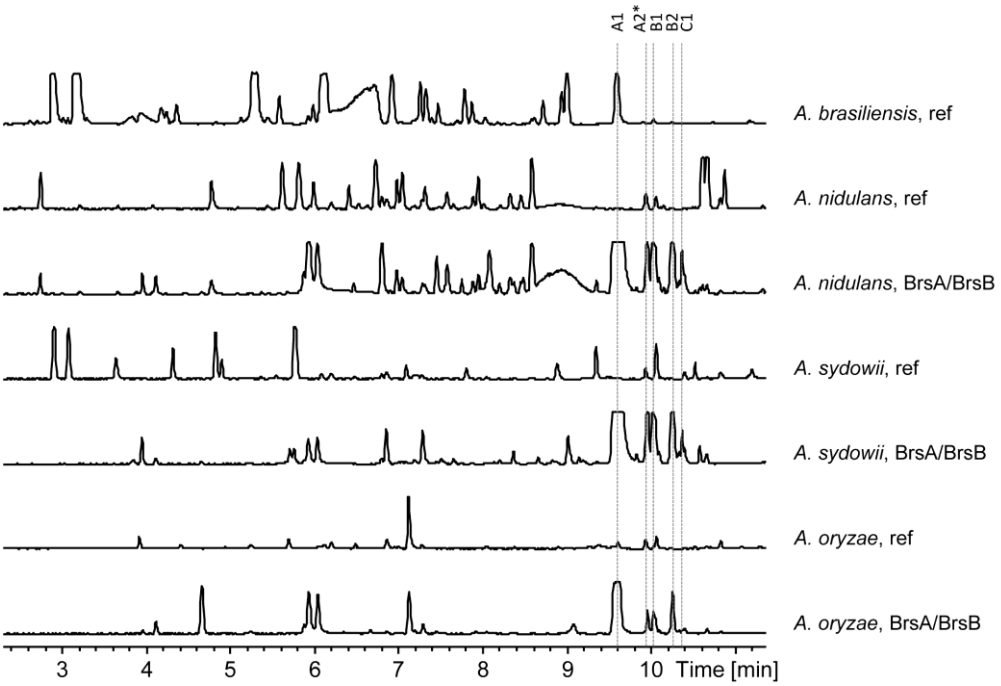


Figure 30. Heterologous expression of the two genes brsA and brsB in the three hosts, A. nidulans, A. sydowii, and A. oryzae. A1, B1, B2 and C1 indicates the isolated brasenol species, and A2 indicates a putative brasenol species similar to brasenol A1, but lacking the double bond between C2' and C3'.*

By heterologously expressing the two *brs* genes, it was possible to isolate an additional three analogues (Figure 31). From NMR analysis, the three analogues were found to be constructed from the same twelve-carbon PK chain as brasenol A1, but differ by varying length and saturation of the 3-hydroxy-fatty acid part, further strengthening our hypothesis that the fatty acid part of the molecule was supplied by the primary metabolism of the host.

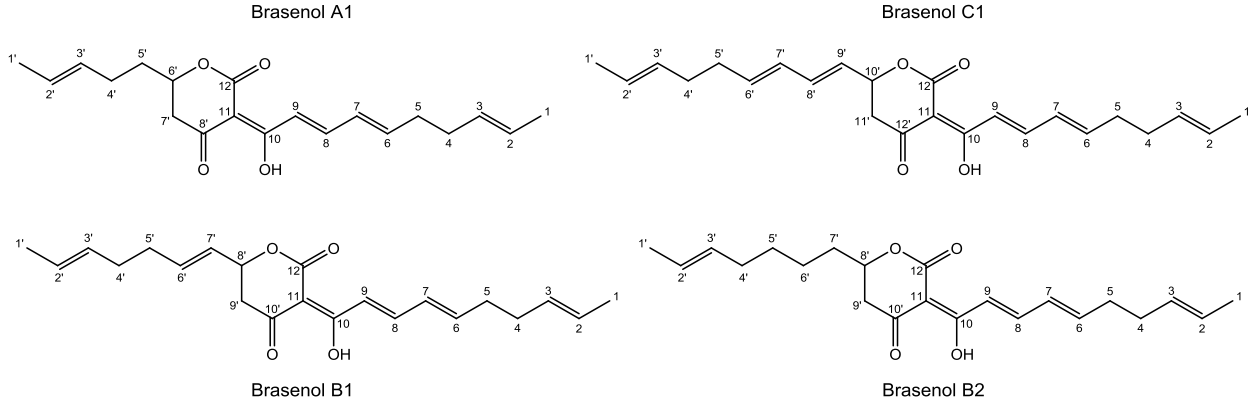


Figure 31. Chemical structures of the four brasenol analogues characterised in this project.

Based on the results from the deletions, heterologous expressions, and NMR studies, a biosynthetic pathway could be suggested (Figure 32): First, the twelve-carbon PK is synthesised by BrsA, most likely from one acetyl-CoA and five malonyl-CoA units. Next, the PK is fused to a 3-hydroxy fatty acid, catalysed by the hydrolase

BrsB, and finally the lactone is formed by esterification between the hydroxyl group and the thioester. This is believed to be catalysed by the esterase BrsC, and also act as a release mechanism. BLAST analysis of BrsC, revealed this gene to have homologues in *A. nidulans*, *A. sydowii*, and *A. oryzae*, explaining why brasenols are still produced in in the three hosts when only BrsA and BrsB are present.

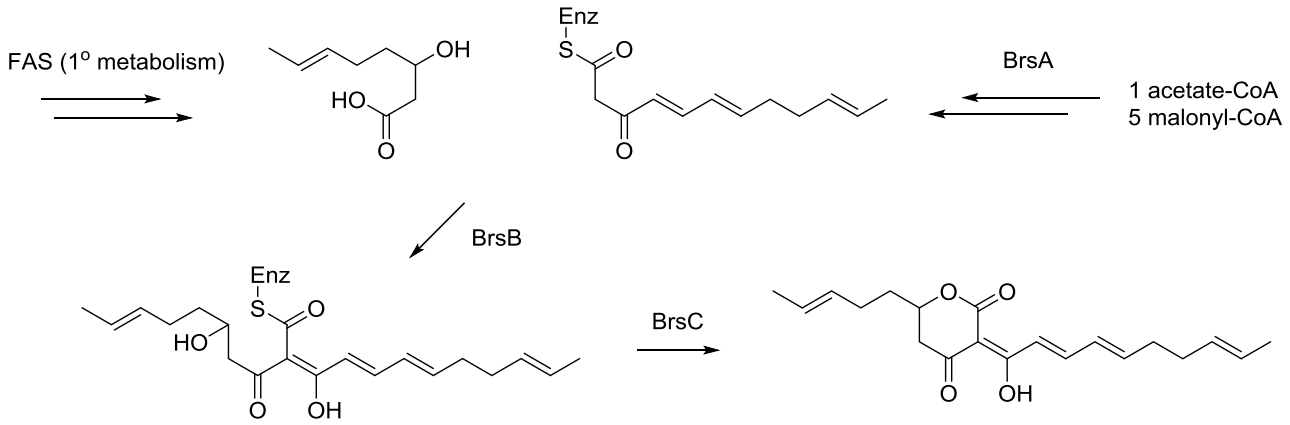


Figure 32. Suggested brasenol biosynthesis based on the deletion the three genes BrsA, BrsB and BrsC, and heterologous expression of BrsA and BrsB, exemplified with brasenol A1.

When BrsA and BrsB were heterologously expressed, sporulation in all three hosts was severely hindered, as depicted in Figure 33. Based on the hypothesis that the 3-hydroxy fatty acids in the brasenol biosynthesis originates from the primary metabolism, lack of regulation when heterologously expressed could be one reason for the observed phenotypes.

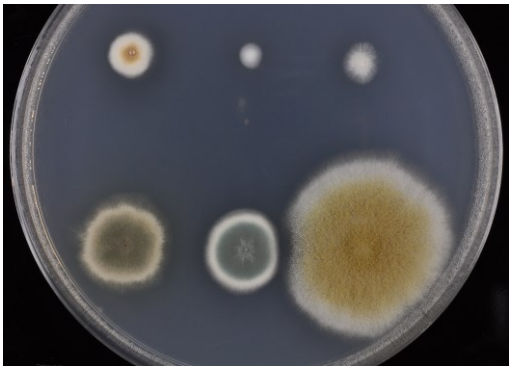


Figure 33. Picture of the three heterologous hosts showing loss of sporulation. Bottom row: Wildtypes of A. nidulans, A. sydowii and A. oryzae (left to right). Top row: brsA/brsB overexpression mutants of A. nidulans, A. sydowii and A. oryzae (left to right).

A certain group of lipids found in most organisms, called oxylipins, have been shown to play a vital part in various stages of regulation and development[10–14], and if the foreign genes work without any regulation, they might interfere with the normal oxylipin metabolism in the hosts. Whether this is the exact cause of the observed effects on sporulation in the three heterologous hosts, is not clear based on this study, why more research into the field is necessary to get a more thorough understanding of the underlying mechanisms.

References

1. Gao, J., Yang, S. & Qin, J. Azaphilones: Chemistry and Biology. *Chem. Rev.* **113,** 4755–811 (2013).

2. Lin, T. F. *et al.* Formation of water-soluble Monascus red pigments by biological and semi-synthetic processes. *J. Ind. Microbiol.* **9,** 173–179 (1992).

3. Kojima, R. *et al.* The relationship between the violet pigment PP-V production and intracellular ammonium level in Penicillium purpurogenum. *AMB Express* **6,** 43 (2016).

4. Arai, T. *et al.* Importance of the ammonia assimilation by Penicillium purpurogenum in amino derivative Monascus pigment, PP-V, production. *AMB Express* **3,** 19 (2013).

5. Tseng, Y. Y., Chen, M. T. & Lin, C. F. Growth, pigment production and protease activity of Monascus purpureus as affected by salt, sodium nitrite, polyphosphate and various sugars. *J. Appl. Microbiol.* **88,** 31–7 (2000).

6. Lin, T. F. & Demain, A. L. Effect of nutrition of Monascus sp. on formation of red pigments. *Appl Microbiol Biotechnol* **36,** 70–75 (1991).

7. Quang, D. N. *et al.* Sassafrins A–D, new antimicrobial azaphilones from the fungus Creosphaeria sassafras. *Tetrahedron* **61,** 1743–1748 (2005).

8. Stierle, A. A. *et al.* Azaphilones from an Acid Mine Extremophile Strain of a Pleurostomophora sp. *J. Nat. Prod.* **78,** 2917–2923 (2015).

9. Reif, B. *et al.* ADEQUATE, a New Set of Experiments to Determine the Constitution of Small Molecules at Natural Abundance. *J. Magn. Reson. Ser. A* **118,** 282–285 (1996).

10. Tsitsigiannis, D. I. & Keller, N. P. Oxylipins as developmental and host-fungal communication signals. *Trends Microbiol.* **15,** 109–118 (2007).

11. Brodhun, F. & Feussner, I. Oxylipins in fungi. *FEBS J.* **278,** 1047–1063 (2011).

12. Kock, J. L. F., Strauss, C. J., Pohl, C. H. & Nigam, S. The distribution of 3-hydroxy oxylipins in fungi. *Prostaglandins Other Lipid Mediat.* **71,** 85–96 (2003).

13. Gessler, N. N. *et al.* Oxylipins and oxylipin synthesis pathways in fungi. *Appl. Biochem. Microbiol.* **53,** 628–639 (2017).

14. Kock, J. L. F. *et al.* A novel oxylipin-associated 'ghosting' phenomenon in yeast flocculation. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* **77,** 401–406 (2000).

Synthetic natural products

Summary of:

Appendix 6: Genes linked to production of secondary metabolites in *Talaromyces atroroseus* revealed using CRISPR-Cas9

Appendix 7: Linker flexibility facilitates module exchange in fungal hybrid PKS-NRPS engineering

Appendix 8: Transforming the lovastatin producing PKS, LovB, into a PKS-NRPS hybrid

Genetic engineering for production of synthetic natural products (Appendix 6+7)

Implementation of new microbiological techniques, such as CRISPR-Cas9, have been an important factor for the possibility of manipulation of the genetic material in species where such endeavours were previously very difficult or not possible with the tools available at the time. In this project genetic engineering of microorganisms was undertaken with various goals in mind. The pigment producing fungus *Talaromyces atroroseus*, described in Chapter 2 (Appendix 1 and 2), was investigated for the potential for genetic modification with the prospect of increasing the pigment yields (Appendix 6). Using the newly described genome editing technique CRISPR-Cas9[1] adapted for use in filamentous fungi[2], it was possible to genetically engineer the strain and disrupt the production of the polyketide-non-ribosomal peptide (PK-NPR) hybrid compounds talaroconvolutin A and B. By establishing a protocol for engineering of *T. atroroseus*, the doors were opened for future manipulation of the strain in order to engineer the pigment production pathways.

Techniques such as CRISPR-Cas9 or USER cloning[3] can also be used to alter the type of product obtained from an organism, rather than the amount. The compound cytochalasin E, produced by *Aspergillus clavatus*, is a PK-NRP hybrid with angiogenic[2] properties[4]. The gene cluster responsible for cytochalasin production was determined in 2011[5] and has in this project been used as a model cluster for biocombinatorial engineering of PK-NRP hybrids (Appendix 7). The cytochalasin gene cluster (Figure 34) consists of a total of eight genes, with a polyketide synthase-non-ribosomal peptide synthetase (PKS-NRPS) and a trans acting enoyl reductase (ccsA and ccsC), responsible for synthesising the backbone of the compound. In addition, a transcription factor and several tailoring enzymes, such as oxidases and a hydrolase are encoded in the gene cluster.

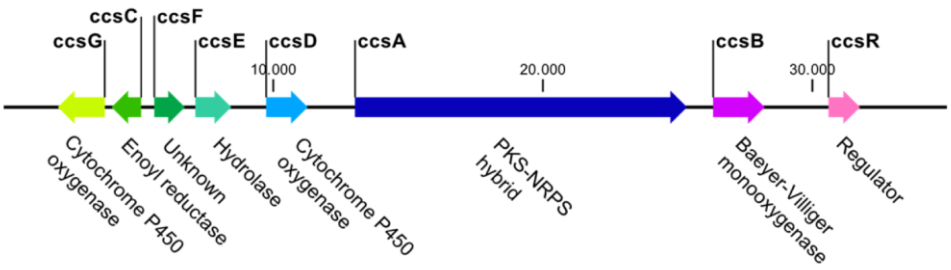


Figure 34. The cytochalasin E gene cluster.

By heterologous expression of the PKS-NRPS gene (*ccsA*) along with the trans acting enoyl reductase (*ccsC*) in *A. nidulans*, HPLC-MS analysis revealed the appearance of a new secondary metabolite. The mass of the compound matched the expected PK-NRP, however with an additional double bond. By purification and NMR

[2] Angiogenesis is the formation of new blood vessels from existing ones, and can be used as a therapeutic target in certain vascular diseases or to prevent proliferation of tumours.[14]

analysis the structure was determined to indeed be a PK-NRP hybrid, albeit not with the structure expected for a cytochalasin derived compound. The compound, named niduclavin, was determined to be composed of the same 16-carbon polyketide chain as the cytochalasins and also linked to a phenylalanine via a tetramic acid (Figure 35). However, the cyclisation of the polyketide, which is normally a tricyclic isoindolone system, was determined to instead have formed a decalin ring system. Furthermore, the α/β-positions in the phenylalanine had been transformed into a double bond.

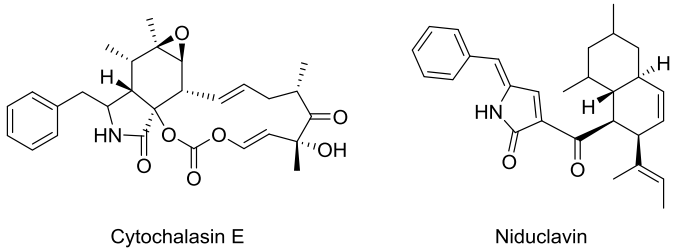


Figure 35. Chemical structures of cytochalasin E and niduclavin.

Even though the compound obtained from the ccsA/ccsC expressing strain was not the expected, it served as a good starting point for further investigations into biocombinatorial approaches. Using bioinformatics, a *ccsA* homologue, *syn2*, was identified in the rice blast fungus *Magnaporthe oryzae*, of which the product was not known but had been suggested to be cytochalasin-like based on gene similarities[6]. The same strategy as for *ccsA* was used, and *syn2* was heterologously expressed along with the trans acting enoyl reductase, also found in the *syn2* gene cluster, *rap2*. Once again a metabolite not normally produced by *A. nidulans* was detected in the extract of the mutant strain. Upon isolation and NMR structural elucidation of the new compound, named niduporthin (Figure 36), it was determined to be structurally similar to niduclavin, but differing in the amino acid part and the methylation pattern of the polyketide. Instead of phenylalanine, the amino acid in niduporthin was a tryptophan, and the compound only had a single methylation in the PK part, rather than the observed three for niduclavin.

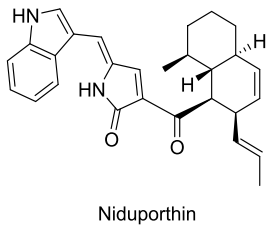
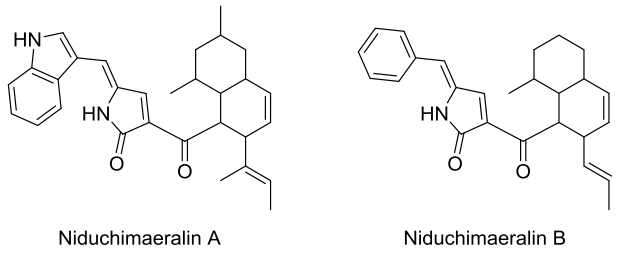


Figure 36. Chemical structure of niduporthin.

Based on the successful results with expression the *ccsA*/*ccsC* and *syn2*/*rap2* gene combinations, as well as additional studies of the variability of the linker between the PKS and NRPS modules in *ccsA*, module swapping

of the PKS-NRPS parts was attempted. Several constructs of the *ccsA* PK and *syn2* NRPS modules were created in order to investigate the optimal location for fusing of the two. This ultimately lead to production of a compound with a mass matching the tri-methylated octaketide part from niduclavin linked to the tryptophan moiety in niduporthin. The structure was confirmed by MS/MS and MS/MS/MS experiments, by comparison with the fragmentation patterns of niduclavin and niduporthin. To test the inverse module swap, the PK part of *syn2* was fused with the NRP part of *ccsA*, once again resulting in a compound matching the expected combination of PK and NRP. The two compounds were named niduchimaeralin A and B and are shown in Figure 37.



Niduchimaeralin A Niduchimaeralin B

Figure 37. Chemical structures of niduchinaeralin A and B.

The observed structures of the four compounds did not match the expected, as a different cyclisation pattern resulted in compounds more similar to PKS-NRPS products such as talaroconvolutin A[7], myceliotermophin E[8] and equisitin[9]. One possible explanation could be the double bond introduced in the amino acid part of the compounds, altering the reactivity of the dienophile undergoing a [4+2] cycloaddition. The origin of the double bond is not clear, but it is speculated to be a result of endogenous activities in *A. nidulans*. A gene encoding a dioxygenase which was thought to be involved in the unwanted double bond formation was deleted, but this did not result in any changes to the products, and heterologous expression of *ccsA* and *ccsC* in *A. niger* resulted in the same products as in *A. nidulans*. From a chemical perspective, adding an additional conjugation is expected to shift the activation to the double bond next to the ketone, thereby promoting ring formation at this position (Figure 38B) rather than at the double bond of the lactam/tetramic acid part, as illustrated in Figure 38A.

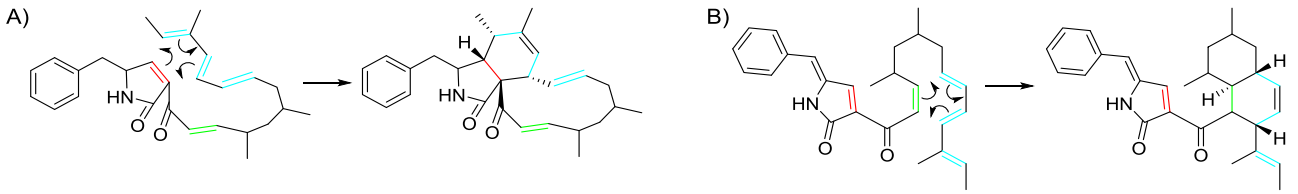


Figure 38. Proposed cyclisation patterns for A) a cytochalasin E precursor and B) for niduclavin. The red bond corresponds to the double bond taking part in the [4+2] cycloaddition in cytochalasin biosynthesis, and the green bond corresponds to the one double bond undergoing the cyclisation in in niduclavin and related compounds.

Transforming the lovastatin producing PKS, LovB, into a PKS-NRPS hybrid (Appendix 8)

The polyketide synthase LovB is responsible for production of the cholesterol-lowering compound lovastatin and its precursor dihydromonacolin L acid in *A. terreus* (Figure 39). This well-studied PKS holds an inactive condensation (C) domain of which the function have been speculated upon by several researchers.[10–12] Usually C domains are not found in PKS's but rather in NRPS's and PKS-NRPS hybrids, and it has therefore been suggested that LovB has evolved from a PKS-NRPS hybrid, but has lost its NPRS activity. In this project, attempts at engineering LovB with the aim of obtaining a functional PKS-NPRS was done using the NPRS module from the PKS-NRPS *ccsA* from *A. clavatus* (Appendix 8).

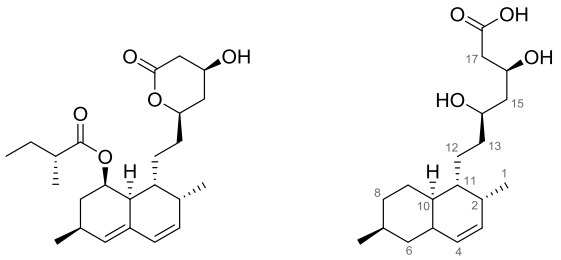


Figure 39. Structures of lovastatin and dihydromonacolin L acid.

Engineering of two versions of LovB/CcsA hybrids and heterologous expression of these in *A. nidulans*, along with the trans-acting enoyl reductase LovC resulted in two mutants, Mutant 1 and Mutant 2 (Figure 40).

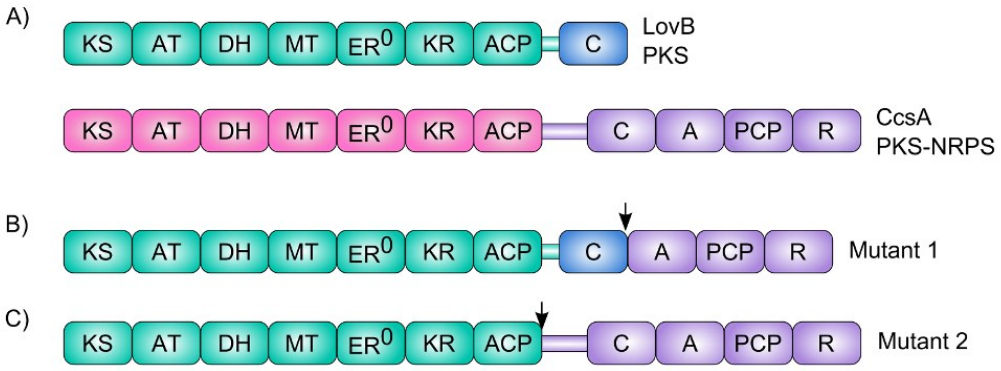


Figure 40. A) The LovB PKS from A. terreus, and the CcsA PKS-NRPS from A. clavatus. B) Mutant 1 was constructed by extending the existing LovB PKS with the A, PCP, and R domains from CcsA. C) Mutant 2 was constructed by replacing the C domain of LovB with the entire NRPS module from CcsA. Arrows indicate points of fusion.

Analysis by LC-MS of the secondary metabolite profile of Mutant 1 revealed production of the lovastatin precursor dihydromonacolin L acid (Figure 39), a result also reported by other research groups[13], but no production of any new nitrogen-containing products. Similarly, analysis of Mutant 2, however, excitingly revealed the production of a new possible PKS-NRPS product with a m/z of 388.2278, corresponding to a molecular formula of $C_{26}H_{30}NO_2$.

Structural analysis by NMR revealed the compound to indeed be a PK-NRP hybrid compound similar to the compounds obtained from the *ccsA* and *syn2* PKS-NRPSs (see Appendix 7), but consisting of a linear octaketide part attached to phenylalanine via a tetramic acid moiety. The compound was named terreclavin and is depicted in Figure 41.

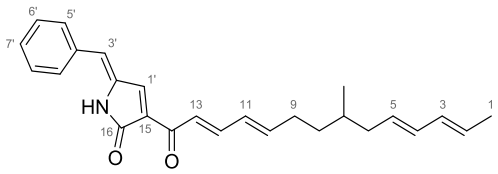


Figure 41. Structure of terreclavin.

Several features, aside from the addition of phenylalanine, separates terreclavin from the normal LovB product lovatstain. First of all the PK chain is constructed from only 16 carbons, and is therefore a single extension shorter than the lovastatin PK. Secondly, the product is linear, despite the fact that the amount and location of double bonds in the PK could allow for a [4+2] cycloaddition between C-2, C-5, C-10, and C-11. Conformational and structural obstacles, however, are expected to prevent the cyclisation from happening, as the shorter and less reduced PK chain is expected to not be able to undergo the same folding as easily as during normal lovastatin biosynthesis. Furthermore, the amount of reductions of the octaketide chain are significantly fewer than in the native products (see Figure 39), with double bonds between C-10 and C-11, and

C-12 and C-13, and no reduction at all at C-14. In comparison, both lovastatin and dihydromonacolin L acid are fully reduced at C-10, and reduced to the hydroxy group at C-14. A common feature for the three compounds, however, is the methylation at C-7. **Error! Reference source not found.** shows the proposed biosynthesis of terreclavin, including formation of the PK-NRP backbone as well as cyclisation and formation of the central tetramic acid part.

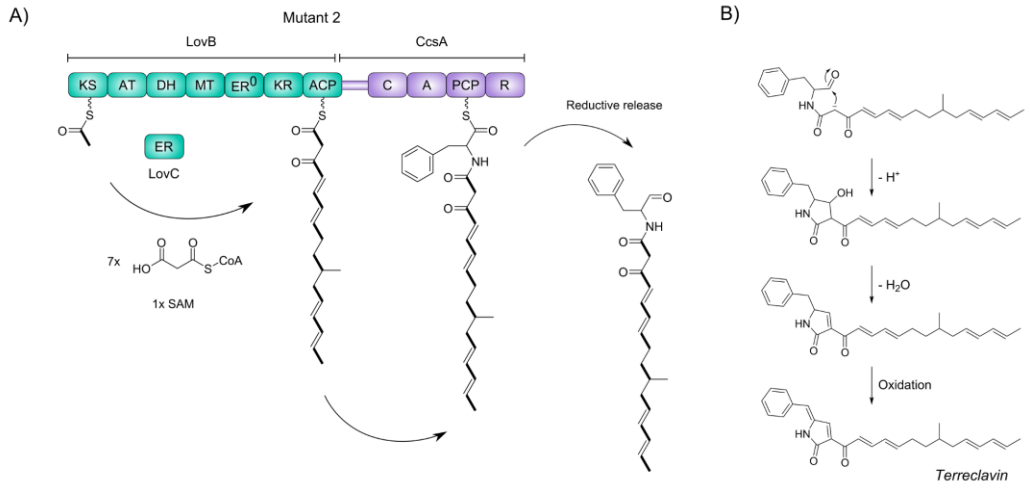


Figure 42. Proposed biosynthesis of terreclavin. A) Construction of the PK-NRP backbone by LovB/LovC. B) Cyclisation and additional tailoring needed to obtain the final product.

Based on the obtained product and previous studies on the LovB PKS and the function of the C domain, the question of its exact role is still not clearly answered. Our results seem to agree with the hypothesis that the C domain play a role in both the cyclisation of the product and possibly in deciding the length of the PK chain. More specifically, our results suggest that the length of the PK chain in affected by the absence or presence of the C domain and that this play a supporting role during lovastatin biosynthesis, as PKSs are normally stringent with regard to the length of their product, and the absence in this study resulted in a PK chain which was one extension shorter than what was expected.

References

1. Reis, A., Hornblower, B., Robb, B. & Tzertzinis, G. CRISPR / Cas9 and Targeted Genome Editing : A New Era in Molecular Biology. *New Engl. BioLabs* 1–4 (2007). at <https://www.neb-online.de/wp-content/uploads/2015/04/NEB_featurearticle_crispr_cas9_summer2014_lowres.pdf>

2. Nødvig, C. S., Nielsen, J. B., Kogle, M. E. & Mortensen, U. H. A CRISPR-Cas9 System for Genetic Engineering of Filamentous Fungi. *PLoS One* **10,** e0133085 (2015).

3. Geu-Flores, F., Nour-Eldin, H. H., Nielsen, M. T. & Halkier, B. a. USER fusion: A rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. *Nucleic Acids Res.* **35,** 0–5 (2007).

4. Udagawa, T. *et al.* Cytochalasin E , an Epoxide Containing Aspergillus-Derived Fungal Metabolite , Inhibits Angiogenesis and Tumor Growth. *J. Pharmacol. Exp. Ther.* **294,** 421–427 (2000).

5. Qiao, K., Chooi, Y. H. & Tang, Y. Identification and engineering of the cytochalasin gene cluster from Aspergillus clavatus NRRL 1. *Metab. Eng.* **13,** 723–732 (2011).

6. Song, Z. *et al.* Heterologous Expression of the Avirulence gene ACE1 from the fungal rice pathogen Magnaporthe oryzae. *Chem. Sci.* **6,** 4837–4845 (2015).

7. Suzuki, S. *et al.* Antifungal Substances Against Pathogenic Fungi, Talaroconvolutins, from Talaromyces convolutus. *J. Nat. Prod.* **63,** 768–772 (2000).

8. Yang, Y. L. *et al.* Cytotoxic polyketides containing tetramic acid moieties isolated from the fungus myceliophthora thermophila: Elucidation of the relationship between cytotoxicity and stereoconfiguration. *Chem. - A Eur. J.* **13,** 6985–6991 (2007).

9. Hazuda, D. *et al.* Isolation and characterization of novel human immunodeficiency virus integrase inhibitors from fungal metabolites. 63–70 (1998).

10. Boettger, D., Bergmann, H., Kuehn, B., Shelest, E. & Hertweck, C. Evolutionary imprint of catalytic domains in fungal PKS-NRPS hybrids. *Chembiochem* **13,** 2363–73 (2012).

11. Ma, S. M. *et al.* Complete reconstitution of a highly reducing iterative polyketide synthase. *Science (80-.).* **326,** 589–592 (2009).

12. Hendrickson, L. *et al.* Lovastatin biosynthesis in Aspergillus terreus: Characterization of blocked mutants, enzyme activities and a multifunctional polyketide synthase gene. *Chem. Biol.* **6,** 429–439 (1999).

13. Kennedy, J. Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. *Science (80-.).* **284,** 1368–1372 (1999).

14. Ferrara, N. & Kerbel, R. S. Angiogenesis as a therapeutic target. *Nature* **438,** 967–974 (2005).

Outlook and Concluding Remarks

This PhD project has dealt with discovery and characterisation of the natural products from microbial sources, more specifically fungal secondary metabolites. Various species have been investigated, both for their usefulness as producers of beneficial compounds, but also for their use as basis for genetic engineering of novel synthetic natural products.

In one study, the filamentous fungus *Talaromyces atroroseus* was found to be the producer of a class of new azaphilones belonging to the class of red *Monascus* pigments. Similarly, *Aspergillus neoglaber* was found to also produce several azaphilone derived compounds, of which two novel yellow and a novel violet pigment was characterised.

A range of different tools have been utilised for obtaining and analysing the compounds of interest, and high-resolution tandem mass spectrometry proved to be a valuable tool for dereplication in conjunction with diode array detection and high-performance liquid chromatography. Additionally, nuclear magnetic resonance spectrometry was essential for structural elucidation of all of the compounds characterised.

With society and the research community entering the genomic era, the use of bioinformatics approaches is becoming increasingly available, and the approach was implemented during this project for identifying the biosynthetic genes responsible for production of a novel *A. brasiliensis* biomarker using comparative genomics. Furthermore, the use of molecular biology tools enabled expression of the biosynthetic genes in several heterologous hosts, revealing an intriguing interaction between the foreign genes and the primary metabolism, by disruption of sporulation in the hosts.

Genetic engineering was moreover used for construction of synthetic natural products, by expression of the cytochalasin-producing PKS-NRPS, CcsA, from *A. clavatus* and the CcsA-homologue Syn2 from the rice blast fungus *Magnaporthe oryzae*, in *A. nidulans*. This resulted in production of two novel PK-NRP products, and by swapping the PKS and NRPS modules between CcsA and Syn2, production of another two novel chimeric products was facilitated, thereby contributing to the future development of synthetic natural products.

Lastly, the lovastatin producing PKS LovB was investigated with the goal of constructing a functional PKS-NRPS. The LovB PKS contains a non-functional condensation domain, and by replacing this with an entire NRPS module, a novel PK-NRP hydrid product was obtained. Based on the structure of the compound the hypothesis that the C domain supports cyclisation in lovastatin biosynthesis was strengthened, and we further hypothesised that it is also influencing the PK chain length in the lovastatin biosynthetic pathway was made.

In summary this thesis has described the characterisation of a range of novel compounds, in addition to a variety of strategies for discovery of natural products, based on classical analytical techniques as well as new bioinformatics and biocombinatorial driven methods.

Appendices

Appendix 1

Atrorosins: a new subgroup of *Monascus* pigments from *Talaromyces atroroseus*

Thomas Isbrandt[1], Gerit Tolborg[1], Mhairi Workman[2], Thomas Ostenfeld Larsen[1*]

[1] Department of Biotechnology and Biomedicine, Søltofts Plads, building 221, 2800 Kongens Lyngby, Denmark

[2] Novo Nordisk A/S, Smørmosevej 10-12, 2880 Bagsværd, Denmark

*Corresponding author

Abstract

A new series of azaphilone pigments named atrorosins have been isolated from the filamentous fungus *Talaromyces atroroseus*. Atrorosins have a similar azaphilone scaffold as the orange *Monascus* pigment PP-O, and are unique by their incorporation of amino acids into the isochromene system. Despite that the atrorosin precursor PP-O, during fermentation was initially produced as two isomers (3:2, *cis:trans* ratio), the atrorosins were surprisingly almost exclusively (99.5%) produced as the *cis* form, indicating that amino acid incorporation into the core polyketide is enzyme catalysed. When grown on complex media a whole range of atrorosins are produced, whereas individual atrorosins can be produced selectively during fermentation by supplementing with the desired amino acid.

Keywords: *Talaromyces*, *azaphilone, Monascus pigments, structural elucidation*

Introduction

Colouring agents, pigments, or dyes, and their use as food additives is a matter that has gained increasing attention during the past 10-15 years. Concerns about the safety of synthetic colorants have made people turn their attention towards natural pigments, such as anthocyanines, carmine and β-carotene. There are however also some challenges regarding these compounds. This includes properties such as stability and solubility of anthocyanines and β-carotene, and allergic reactions to carmine products.[1]

Monascus pigments are a class of natural fungal pigments originally identified in filamentous fungi belonging to the genus *Monascus*, however these compounds have also have been found in species belonging to the genera *Penicillium* and *Talaromyces*. Species within the *Monascus* genus have been used as food preservatives and colouring for several hundred years and are still used in Asia for production of red rice *koji*[2]. However, since production of mycotoxins such as citrinin, and other compounds with undesired effects for food ingredients such as lovastatin (monacolin K), have been associated with *Monascus*,[3] fungi from this genus are not accepted for use on the European and North American market.[4]

Structurally, the *Monascus* pigments represent a subgroup belonging to the class of compounds known as azaphilones, a hugely diverse family of fungal secondary metabolites.[5] Six pigments, first described by Blanc et al[2], are the prototypes for the family of *Monascus* pigments: the yellow monascin and ankaflavin, the orange rubropunctatin and monascorubrin, and the red rubropunctamine and monascorubramine (Figure 1).

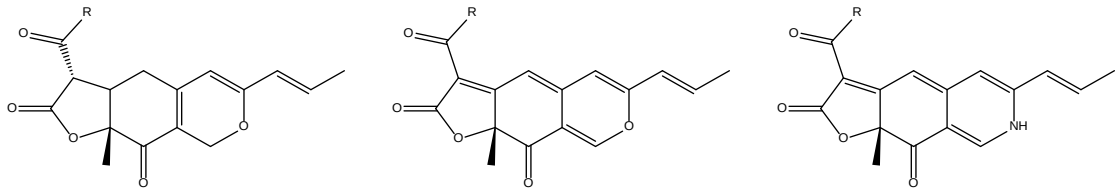


Figure 1. The six original Monascus pigments, with colours spanning from yellow to red.

The filamentous fungus *Talaromyces atroroseus* (previously *Penicillium purpurgenum)*, has recently been shown to secrete large amounts of coloured compounds belonging to the *Monascus* pigments, without production of mycotoxins.[6–10] Some of the pigments are known, such as PP-O, PP-V, PP-Y/monascorubrin, and monascorubramine,[10–13] but a majority of the produced pigments have not been structurally characterised. This study presents the structures of a novel class of *Monascus* pigment, atrorosins, produced by *T. atroroseus*.

Materials and Methods

Strain, Reagents and Instruments

The strain used throughout this study was *Talaromyces atroroseus* IBT 11181 (DTU strain collection). Spores were propagated on CYA agar plates and incubated at 30 °C for 7 days. Spores were harvested with 0.9% sodium chloride solution (NaCl), filtered through mira-cloth, centrifuged and then re-suspended in 0.9% NaCl solution. The spore concentration was determined by using a Burker-Turk counting chamber. All cultivations were inoculated to give an initial spore concentration of 106 spores/ml.

All purchased solvents and reagents were acquired from Sigma-Aldrich (St. Louis, Missouri, USA), ultra-pure water was made with a Milli-Q system (Millipore, Burlington, Massachusetts, USA).

Measuring of optical rotation was done on a Perkin-Elmer 341 Polarimeter (Perkin Elmer, Waltham, Massachusetts, USA) using a 10 cm cell.

Ultra-high Performance Liquid Chromatography-High Resolution Mass Spectrometry (UHPLC-HRMS) was performed on an Agilent Infinity 1290 UHPLC system (Agilent Technologies, Santa Clara, CA, USA) equipped with a diode array detector. Separation was obtained on an Agilent Poroshell 120 phenyl-hexyl column (2.1 × 250 mm, 2.7 µm) with a linear gradient consisting of water (A) and acetonitrile (B) both buffered with 20 mM formic acid, starting at 10% B and increased to 100% in 15 min where it was held for 2 min, returned to 10% in 0.1 min and remaining for 3 min (0.35 mL/min, 60 °C). An injection volume of 1 µL was used. MS detection was performed in positive detection mode on an Agilent 6545 QTOF MS equipped with Agilent Dual Jet Stream electrospray ion source with a drying gas temperature of 250 °C, gas flow of 8 L/min, sheath gas temperature

of 300 °C and flow of 12 L/min. Capillary voltage was set to 4000 V and nozzle voltage to 500 V. Mass spectra were recorded at 10, 20 and 40 eV as centroid data for *m*/*z* 85–1700 in MS mode and *m*/*z* 30–1700 in MS/MS mode, with an acquisition rate of 10 spectra/s. Lock mass solution in 70:30 methanol:water was infused in the second sprayer using an extra LC pump at a flow of 15 µL/min using a 1:100 splitter. The solution contained 1 µM tributylamine (Sigma-Aldrich) and 10 µM Hexakis(2,2,3,3-tetrafluoropropoxy)phosphazene (Apollo Scientific Ltd., Cheshire, UK) as lock masses. The $[M + H]^+$ ions (*m*/*z* 186.2216 and 922.0098 respectively) of both compounds were used.

1D and 2D NMR spectra were recorded on a Bruker Avance 800 MHz spectrometer (Bruker, Billerica, MA, USA) located at the Department of Chemistry at the Technical University of Denmark. NMR spectra were acquired using standard pulse sequences. The solvent used was either DMSO-d_6 or MeOD, which were also used as references with signals at δ_H = 2.50 ppm and δ_C = 39.5 ppm for DMSO-d_6 and δ_H = 3.31 ppm and δ_C = 49.0 ppm for CD_3OD. Data processing and analysis was done using TopSpin 3.5pl7 (Bruker). *J*-couplings are reported in hertz (Hz) and chemical shifts in ppm (δ). For all natural pigments, 1D proton, edHSQC, and HMBC were recorded, and for atrorosin S, LR-HSQMBC, 1,n-ADEQUATE, and 1,1-ADEQUATE were also measured. Compounds were measured in MeOD, except atrorosin Q which was measured in DMSO-d_6. For the semi-synthesised nature identical pigments, only 1D proton and edHSQC were recorded.

Cultivation

The medium for shake flask cultivations was composed of sucrose (7.5 g/L), glucose (0.375 g/L), KH_2PO_4 (10 g/L), NaCl (1 g/L), $MgSO_4 \cdot 7\ H_2O$ (2 g/L), KCl (0.5 g/L), $CaCl_2 \cdot H_2O$ (0.1 g/L) and trace metal solution (2 mL/L). The trace metal solution consisted of $CuSO_4 \cdot 5\ H_2O$ (0.4 g/L), $Na_2B_4O_7 \cdot 10\ H_2O$ (0.04 g/L), $FeSO_4 \cdot 7\ H_2O$ (0.8 g/L), $MnSO_4 \cdot H_2O$ (0.8 g/L), $Na_2MoO_4 \cdot 2\ H_2O$ (0.8 g/L), $ZnSO_4 \cdot 7\ H_2O$ (8 g/L). The nitrogen source was 0.1 M each of an L-amino acid. Controls were done using 0.1 M of KNO_3. The pH of the medium was adjusted to pH 5 with aqueous NaOH and HCl. Cultivations were carried out in non-baffled shake flasks at 30 °C and 150 rpm in rotary shaking incubators (Forma orbital shaker, Thermo Fisher Scientific, US) with a volume of 100 ml. Samples were

taken after 96 hrs. Shake flask experiments were carried out in triplicates. And extractions were done on pooled samples after analysis.

Bioreactor cultivation of PP-O was carried out in duplicates in 1 L bioreactors. The medium for contained sucrose (20 g/L), glucose (1 g/L), KH_2PO_4 (10 g/L), NaCl (1 g/L), $MgSO_4 \cdot 7\ H_2O$ (2 g/L), KCl (0.5 g/L), $CaCl_2 \cdot H2O$ (0.1 g/L) and trace metal solution (2 mL/L). 2 g/L of KNO_3 were used as nitrogen source. The cultivation was stopped after 53 h at which point it was harvested. The cultivation was performed at 30 °C, 800 RPM, 1 VMM and a pH of 4.5.

The bioreactors were Sartorius 1 L bioreactors (Satorius, Stedim Biotech, Goettingen, Germany) with equivalent working volumes, equipped with 2 Rushton six-blade disc turbines. The pH electrode (Mettler Toledo, OH/USA) was calibrated according to manufacturer's standard procedures. The bioreactor was sparged with sterile atmospheric air and off-gas concentrations of oxygen and carbon dioxide were measured with a Prima Pro Process Mass Spectrometer (Thermo-Fischer Scientific, Waltham, MA/USA), calibrated monthly with gas mixtures containing 5 % (v/v) CO_2, 0.04 % (v/v) ethanol and methanol, 1 % (v/v) argon, 5 % (v/v) and 15 % (v/v) oxygen all with nitrogen as carrier gas (Linde Gas, AGA, Enköping, Sweden). The pH was controlled by automatic addition of 2 M NaOH and H_2SO_4.

Purification of PP-O isomers

For purification of PP-O, fermentation liquid was extracted two times with EtOAc to give 1.2 grams of crude extract which was loaded onto an Isolera One (Biotage, Uppsala, Sweden) automated flash system equipped with a diol column and eluted stepwise with dichloromethane (DCM), DCM:EtOAc (1:1), EtOAc, EtOAc:MeOH (1:1), and MeOH. The enriched orange fraction (0.9 g) was subjected to solid phase extraction (SPE) to remove co-eluting impurities. The diol fraction was dissolved in MeOH and loaded onto an equilibrated Oasis WAX solid phase extraction (SPE) column (Waters, Milford, Massachusetts, USA) and washed with 3 column volumes (CVs) of MeOH followed by elution of the orange pigment with 3 CVs of 2% TFA in MeOH. The orange fraction was purified on the same Gilson semi-prep HPLC system as the atrorosins, but using a water/MeOH gradient

with 50 ppm TFA and a Kinetex Core-Shell C18 column (250 mm x 10 mm, 5 µm), in order to separate the two isomers.

Purification of atrorosins

Filtered and centrifuged fermentation broth was extracted three times, with 1/3 volume of EtOAc, at pH 3 (adjusted with formic acid). The combined EtOAc phases were evaporated to 100 mL and extracted twice with water (1:1) at pH 8 (adjusted with ammonium hydroxide). The water phase was readjusted to pH 3 with FA and extracted two time with EtOAc, followed by evaporation, to yield >95% pure pigment (a mixture of atrorosins and N-amino acid monascorubramine, ratio>10:1 by LC-DAD-MS). The final purification was performed on a Gilson 332 semi-prep HPLC system equipped with a Gilson 172 diode array detector, using a LUNA II C18 column (250 mm x 10 mm, 5 µm, Phenomenex), with a water/acetonitrile gradient with 50 ppm trifluoroacetic acid (TFA).

Derivatisation of PP-O

For the chemical derivatisation of PP-O to form nature identical atrorosins, PP-O and a 10-15-times molar excess of the amino acid was dissolved in a 1:1 mixture of MeOH and water with 100 ppm TFA (v/v). The solution was heated to 40 °C for 15 min with stirring, and the excess amino acid and unreacted PP-O was removed on an Evolute Express WAX SPE column (Biotage). For lysine, the two isomers were separated by semi-preparative HPLC using a Waters 600 Controller with a 996 photodiode array detector (Waters, Milford, MA, USA) equipped with a Luna II C18 column (250 × 10 mm, 5 µm, Phenomenex), using a H_2O/acetonitrile gradient with 50 ppm TFA.

Results and discussion

Cultivation of *T. atroroseus* leads to isolation of *trans*-PP-O and a novel azaphilone class, atrorosins.

Liquid culture fermentation of the filamentous fungus *Talaromyces atroroseus*[6] (IBT11181, DTU strain collection) gave a culture liquid with an intense red colour. Upon analysis of the ethyl acetate extract of the fermentation liquid by Ultra High Performance Liquid Chromatography coupled to Diode Array Detection and

High-Resolution tandem Mass Spectrometry (UHPLC-DAD-MS/HRMS), a plethora of different pigmentss were observed. Among them, several known *Monascus* pigments (e.g. PP-R, PP-V, PP-Y/monascorubrin, and monascorubramine) were identified, along with numerous unknown pigment species.

Careful analysis of the data also showed the presence of what appeared to be a 3:2 ration of two distinct isomers of the orange azaphilone PP-O, based on integration of their identical UV spectra. The two isomers were isolated and separated using reverse phase HPLC to generate enough material for structural analysis. From NMR experiments, we were able to identify the two PP-O isomers to be identical, except for having either a *cis*- or a *trans*- form of the double bond between C-2 and C-3. H-2 and H-3 had a coupling constant of *J*=12.8 Hz, establishing the alkene to be in a Z-configuration.[12,17] NMR spectra for the PP-O isomers can be found in Supporting Information S1, and Supporting Information S2 shows NMR shifts (Table S1) and structures (Figure S1). Since the *trans* form of PP-O has not been reported before, we have named this new compound *trans*-PP-O, and will refer to the known isomer as *cis*-PP-O.

Next we turned our interest towards one of the major unknown red compounds with a mass-to-charge ratio (m/z) of 500.1912, and a UV-VIS absorption spectrum similar to that of known red *Monascus* pigments, i.e. with strong absorption at 430 and 520 nm. Figure 2 shows an example of the UV-VIS profile at 520 nm of a shake flask fermentation using KNO_3 as nitrogen source.

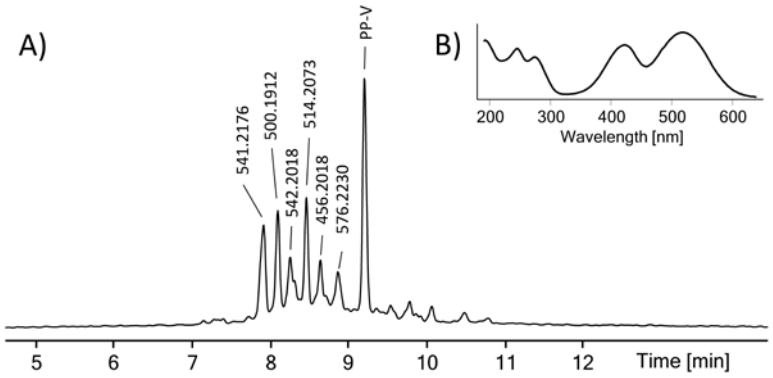


Figure 2. A) UV-VIS chromatogram (520±10nm) of the EtOAc extract from shake flask cultivation of T. atroroseus using KNO3 as nitrogen source. The m/z or identity of the major peaks has been assigned according to the LC-DAD-MS data. B) UV-VIS spectrum of m/z=500.

The unknown red pigment was purified by reverse phase chromatography to give a dark red amorphous solid. From HRMS, a molecular formula of $C_{26}H_{29}N_2O_9$ (DBE=13) was determined (calc. 500.1915, mass accuracy = 0.6 ppm). 1D and 2D NMR was used to determine the structure of the compound. NMR spectra can be found in Supporting Information S1, and chemical shift values are listed in Table S2. ^{1}H-NMR and HSQC revealed five alkene CH-groups, two methyl groups, seven CH_2-groups, and one alkane CH-group. ^{13}C-NMR and HMBC revealed an additional 11 quaternary carbons of which five were carbonyls, five alkene carbons, and one quaternary alkane. Additionally, two carboxylic acids and an amide were present in the molecule (H and C shifts are listed in table 1).

Tabel 1. Proton and carbon shifts for atrorosin S in CD_3OD.

#	δ_H	δ_C
1	-	166.8
2	6.43 (d, 11.8)	130.9
3	6.94 (d, 11.7)	133.8
4	-	149.7
5	6.81 (s)	119.6
6	-	151.8
7	6.72 (s)	117.7
8	-	168.0
9	-	86.8
9-CH3	1.67 (s)	30.1
10	-	195.4
11	-	98.8
12	8.57 (s)	142.8
13	-	174.7
14	-	125.4
15	-	198.5
16	2.82 (m)	40.9
17	1.59 (quin, 7)	25.8
18	1.31 (m)	23.3/30.0/32.5
19	1.31 (m)	23.3/30.0/32.5
20	1.31 (m)	23.3/30.0/32.5
21	1.31 (m)	23.3/30.0/32.5
22	0.89 (t, 6.2)	13.9
1'	-	169.1
2'	5.12 (s)	67.3
3'	4.28 (dd, 12.2, 5.6) / 4.09 (d, 12.2)	62.5

DQF-COSY showed ^{3}J-couplings between H-2' and H-3', as well as between H-2 and H-3. H-2 and H-3 had a coupling constant of J=11.8 Hz, suggesting the alkene to be in a Z-configuration. Furthermore, an aliphatic chain consisting of H-16 to H-22 could be identified.

HMBC provided long-range H-C-couplings and could link H-16 and H-17 to the carbonyl C-15. H-7 also showed a coupling to C-15, in addition to couplings to C-6, C-8, C-9, C-11, C13, C14 and C-15, while H-12 had couplings to C-6, C-10 and C-11. H-2 and H-5 both exhibited couplings to C-4. Correlations from H-'2 to C-4 and C-12 was also observed in both HMBC and LR-HSQMBC. Correlations between H-7 and C-8 was determined from 1,1-ADEQUATE.

Based on the structural information obtained from the NMR data, an isoquinoline system could be assembled as the core of the molecule, consisting of C-4 to C-12, with C-13 and C-14 forming a five-membered lactone with C-8 and C-9. Attached to the central ring structure, an α,β-unsaturated carboxylic acid and a methyl group could be linked to C-4 and C9, respectively. A heptanyl chain could be linked to C-14 via the ketone C-15. Finally, a serine moiety was determined to be attached through the isoquinoline nitrogen.

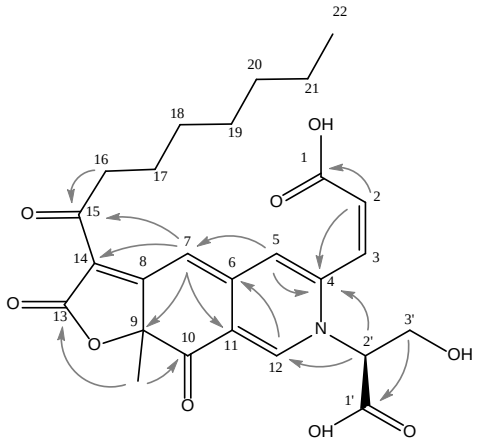


Figure 3. Numbered structure of atrorosin S with key HMBC correlations.

In summary, the compound (Figure 3) is structurally very similar to the *Monascus* pigments PP-O and PP-V[12], but with the incorporation of a serine moiety into the isochromene/isoquinoline system. We have named the

compound atrorosin S, using the one-letter amino acid abbreviation to denote which nitrogen-containing compound has been incorporated.

Amino acid feeding expands and focuses the atrorosin catalogue

The discovery of this new azaphilone, stirred our curiosity as to whether *T. atroroseus* was also able to incorporate other amino acids or nitrogen containing compounds into the azaphilone core, as this had previously been described for monascorubrin[14–16]. The LC-MS data from the crude extract was re-examined and several compounds, matching tentative atrorosins were identified. All amino acid derived atrorosins except arginine, histidine, lysine, methionine and proline could be tentatively identified in the extract.

In order to confirm the identities of the pigments, the fungus was cultivated in shake flasks using each of the remaining 19 amino acids as the sole nitrogen source. The ethyl acetate extract of the filtered culture liquid was analysed by UHPLC-DAD-MS/HRMS for each of the 19 cultivations. For proline, no production of the corresponding atrorosins was neither expected nor detected, as the nitrogen atom in the amino acid is a secondary amine. For each of the remaining amino acids, with the exception of tryptophan, the expected atrorosin was to some degree detected in the extract, as illustrated in Figure 4.

Interestingly, upon cultivation with a single amino acid as the sole nitrogen source, the secondary metabolite profile for nearly all experiments changed to contain primarily a single red compound, namely the atrorosin corresponding to the fed amino acid. For a few of the experiments (glycine, cysteine, threonine and tyrosine) other coloured compounds were also detected, as shown in Figure 4, but still far less when compared to shake flask experiments using KNO_3 as nitrogen source (Figure 2).

For glycine and tryptophan, the pigment production was either very low or almost completely absent. In the chromatogram from the glycine extract, some peaks corresponding to unknown red compounds were present, but the one belonging to the tentative atrorosin G was found only in very small amounts, while no red pigments were found in the tryptophan extract. When grown with tyrosine as nitrogen source, in addition to the expected atrorosin Y, significant amounts of both PP-Y and PP-O were also found in the extract.

The lack of pH-control in the shake flasks was most likely the reason why not all of them were equally successful. Growth experiments in bioreactors with well-defined growth conditions has shown that the optimal pH for pigment production was in the quite narrow range of 4-5 (Tolborg *et al.*, in preparation).

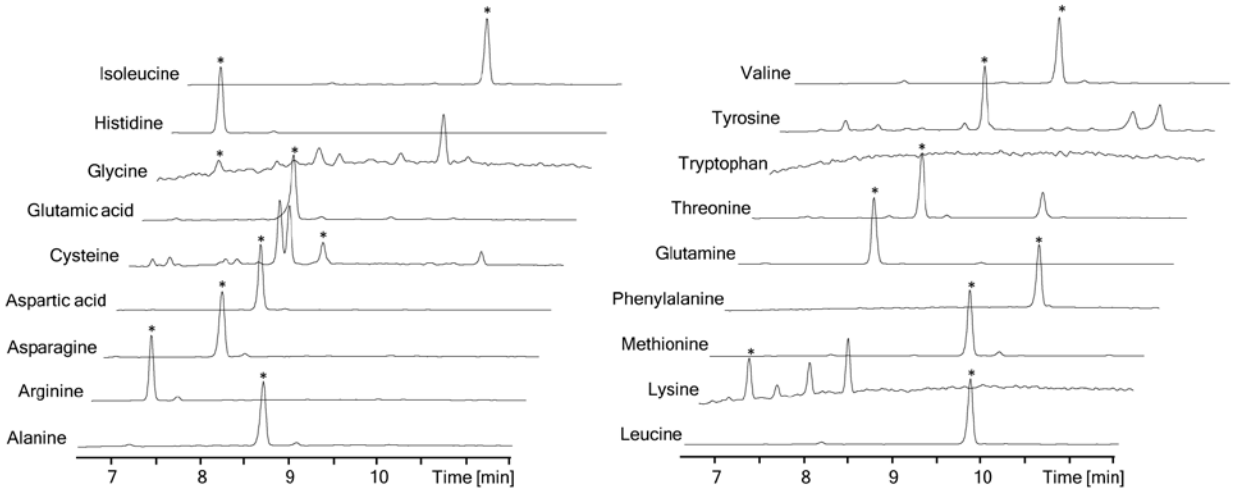


*Figure 4. UV-VIS chromatograms (520±10 nm) of EtOAc extracts from shake flask cultivations of T. atroroseus using 18 different amino acids as sole nitrogen source. * indicates the expected atrorosins, confirmed by HRMS.*

Because of the poorly controlled conditions in the shake flask cultivations, we were only capable of isolating atrorosins from seven of these (D, E, H, L, M, Q and T) in addition to the original atrorosin S. Rather than carrying out bioreactor cultivations with each of the remaining amino acids, a single four litre fermentation, was done and harvested at the time where the PP-O concentration was at its highest, with the aim of derivatising PP-O to produce all nature identical atrorosins.

Semi-synthesis of nature identical atrorosins from *cis*-PP-O

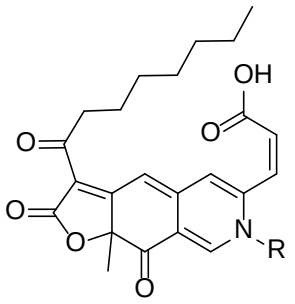
During the bioreactor cultivation, L-amino acids were added in order to produce single atrorosins, and it was therefore expected that these would also contain the L-amino acid upon excretion. However, to ensure that no isomerisation had taken place, we synthesised nature identical atrorosins from *cis*-PP-O and both L- and D-amino acids to compare these against natural pigments isolated from the fermentation broth.

The chemical incorporation of amine containing compounds into the backbone of the *Monascus* pigment monascorubrin/PP-Y has previously been reported[15] and we have shown that the same is possible with PP-O

under basic conditions, however with poor selectivity as several isomers were formed (Tolborg *et al., in preparation*).

In the present study, we learned that though slower, the acid catalysed reaction provides better stereoselectivity during the *in vitro* incorporation of e.g. amino acids, compared to the base catalysed reaction (data not shown). By reacting *cis*-PP-O with amino acids in a 1:1 mixture of water and methanol acidified with 100 ppm triflouroacetic acid we could almost quantitatively obtain each of the desired atrorosins (HPLC data in Supporting Information S4, Figure S2). ^{1}H-NMR and HSQC were recorded for each of the synthesised compounds, and could together with accurate mass and fragmentation patterns from HR-MS/MS confirm formation of the expected products (NMR spectra and chemical shifts for each of the semi-synthesised atrorsins can be found in the Supporting Information S1 and S3). For atrorosin E and S we also incorporated the D-amino acid into *cis*-PP-O to compare the optical rotation with the natural compound in order to confirm whether the amino acid is modified before being excreted by the fungus (values listed in Supporting Information S5). These experiments confirmed our hypothesis, that the amino acids in the natural atrorosins are indeed the more naturally abundant L-amino acids and that the natural atrorosins all contain L-amino acids, when these are added during cultivation.

The remaining atrorosins not isolated from the shake flask cultivations were also made from *cis*-PP-O, and ^{1}H-NMR and HSQC spectra were obtained for these. NMR spectra, and proton and carbon shifts for all characterised atrorosins as well as both isomers of PP-V can be found in Supporting Information S1 and S3, and their structures (except atrorosin S) are shown in Figure 5.



R= Ala: Atrorosin A, Cys: Atrorosin C, Asp: Atrorosin D, Glu: Atrorosin E, Phe: Atrorosin F, Gly: Atrorosin G, His: Atrorosin H, Iso: Atrorosin I, Lys: Atrorosin K, Leu: Atrorosin L, Met: Atrorosin M, Asn: Atrorosin N, Gln: Atrorosin Q, Arg: Atrorosin R, Thr: Atrorosin T, Val: Atrorosin V, Trp: Atrorosin W, Tyr: Atrorosin Y

Figure 5. Chemical structure of 18 atrorosins.

In summary, the filamentous fungus *T. atroroseus* has been found to excrete high amount of a new class of red azaphilone pigments, atrorosins. These are all derivatives of the orange pigment *cis*-PP-O, differing only by their incorporation of L-amino acids, which can either be naturally incorporated during standard growth conditions to give a range of different compounds, or be selectively incorporated by feeding single amino acids during cultivation to give only one version of the pigment. On the contrary to PP-O, that was showed to be naturally produced in both a *cis*- and *trans* form, we find it highly surprising that the atrorosins were almost exclusively (>99.5%) produced as the *cis* isomer. We therefore propose that amino acid incorporation into the core polyketide skeleton in *T. atroroseus* might be enzyme assisted, something that to the best of our knowledge has not previously been suggested, and could even involve transformation of the C-2/C-3 alkene in *trans*-PP-O into the *cis* form seen in atrorosins, since neither *trans*-PP-O nor "*trans*-atrorosins" could be detected after incorporation of amino acids.

Acknowledgements

The authors would like to acknowledge Kasper Enemark-Rasmussen at the Department of Chemistry, Technical University of Denmark, for running NMR experiments.

References

(1) Rohrig, B. *ChemMatters* **2015**, No. October/November, 5.

(2) Blanc, P. J.; Loret, M. O.; Santerre, a L.; Pareilleux, A.; Prome, D.; Prome, J. C.; Laussac, J. P.; Goma, G. *J. Food Sci.* **1994**, *59* (4), 862.

(3) Yang, J.; Chen, Q.; Wang, W.; Hu, J.; Hu, C. *J. Biosci. Bioeng.* 2015, *119* (5), 564.

(4) Eisenbrand, G. *Mol. Nutr. Food Res.* 2006, *50* (3), 322.

(5) Gao, J.; Yang, S.; Qin, J. *Chem. Rev.* 2013, *113* (7), 4755.

(6) Frisvad, J. C.; Yilmaz, N.; Thrane, U.; Rasmussen, K. B.; Houbraken, J.; Samson, R. *PLoS One* 2013, *8* (12), e84102.

(7) Mapari, S. Prodcution of Monascus-like azaphilone pigments. EP 2262862 B1, 2007.

(8) Mapari, S. a S.; Nielsen, K. F.; Larsen, T. O.; Frisvad, J. C.; Meyer, A. S.; Thrane, U. *Curr. Opin. Biotechnol.* 2005, *16* (2), 231.

(9) Mapari, S. a S.; Meyer, A. S.; Thrane, U. *J. Agric. Food Chem.* 2006, *54* (19), 7027.

(10) Ogihara, J.; Kato, J.; Oishi, K.; Fujimoto, Y.; Eguchi, T. *J. Biosci. Bioeng.* 2000, *90* (5), 549.

(11) Ogihara, J.; Kato, J.; Oishi, K.; Fujimoto, Y. *J. Biosci. Bioeng.* 2001, *91* (1), 44.

(12) Arai, T.; Kojima, R.; Motegi, Y.; Kato, J.; Kasumi, T.; Ogihara, J. *Fungal Biol.* 2015, *119* (12), 1226.

(13) Ogihara, J. 2001, *93* (1), 54.

(14) Hajjaj, H.; Klaebe, A.; Loret, M. O.; Tzedakis, T.; Goma, G.; Blanc, P. J. *Appl. Environ. Microbiol.* 1997, *63* (7), 2671.

(15) Lin, T. F.; Yakushijin, K.; Bt, G. H.; Demain, A. L.; Buchi, G. H.; Demain, A. L.; Bt, G. H.; Demain, A. L. *J. Ind. Microbiol.* 1992, *9* (3–4), 173.

(16) Jung, H.; Kim, C.; Kim, K.; Shin, C. S. *J. Agric. Food Chem.* 2003, *51* (5), 1302.

(17) Ogihara, J.; Oishi, K. *J. Biosci. Bioeng.* 2002, *93* (1), 54.

SUPPORTING INFORMATION

Atrorosins: a new subgroup of *Monascus* pigments from *Talaromyces atroroseus*

Thomas Isbrandt[1], Gerit Tolborg[1], Mhairi Workman[2], Thomas Ostenfeld Larsen[1*]

[1] Department of Biotechnology and Biomedicine, Søltofts Plads, building 221, 2800 Kongens Lyngby, Denmark
[2] Novo Nordisk A/S, Smørmosevej 10-12, 2880 Bagsværd, Denmark

*Corresponding author

Contents

S1. NMR spectra for all atrorosins and *cis*- and *trans*-PP-O

The following pages contains NMR spectra for the following compounds:

Atrorosin A

Atrorosin C

Atrorosin D

Atrorosin E

Atrorosin F

Atrorosin G

Atrorosin H

Atrorosin I

Atrorosin K1

Atrorosin K2

Atrorosin L

Atrorosin M

Atrorosin N

Atrorosin Q

Atrorosin R

Atrorosin S

Atrorosin T

Atrorosin V

Atrorosin W

Atrorosin Y

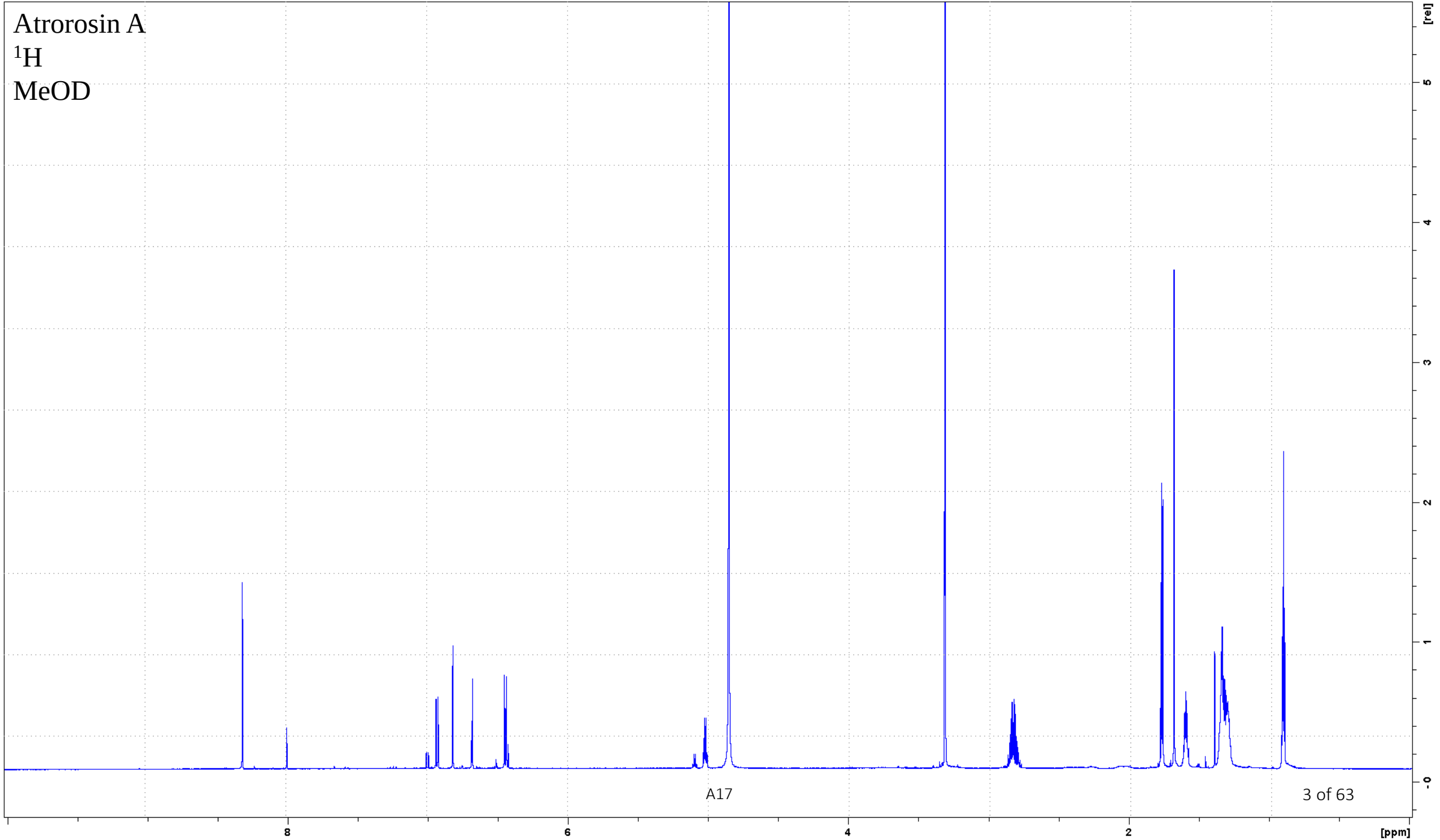
cis-PP-O

trans-PP-O

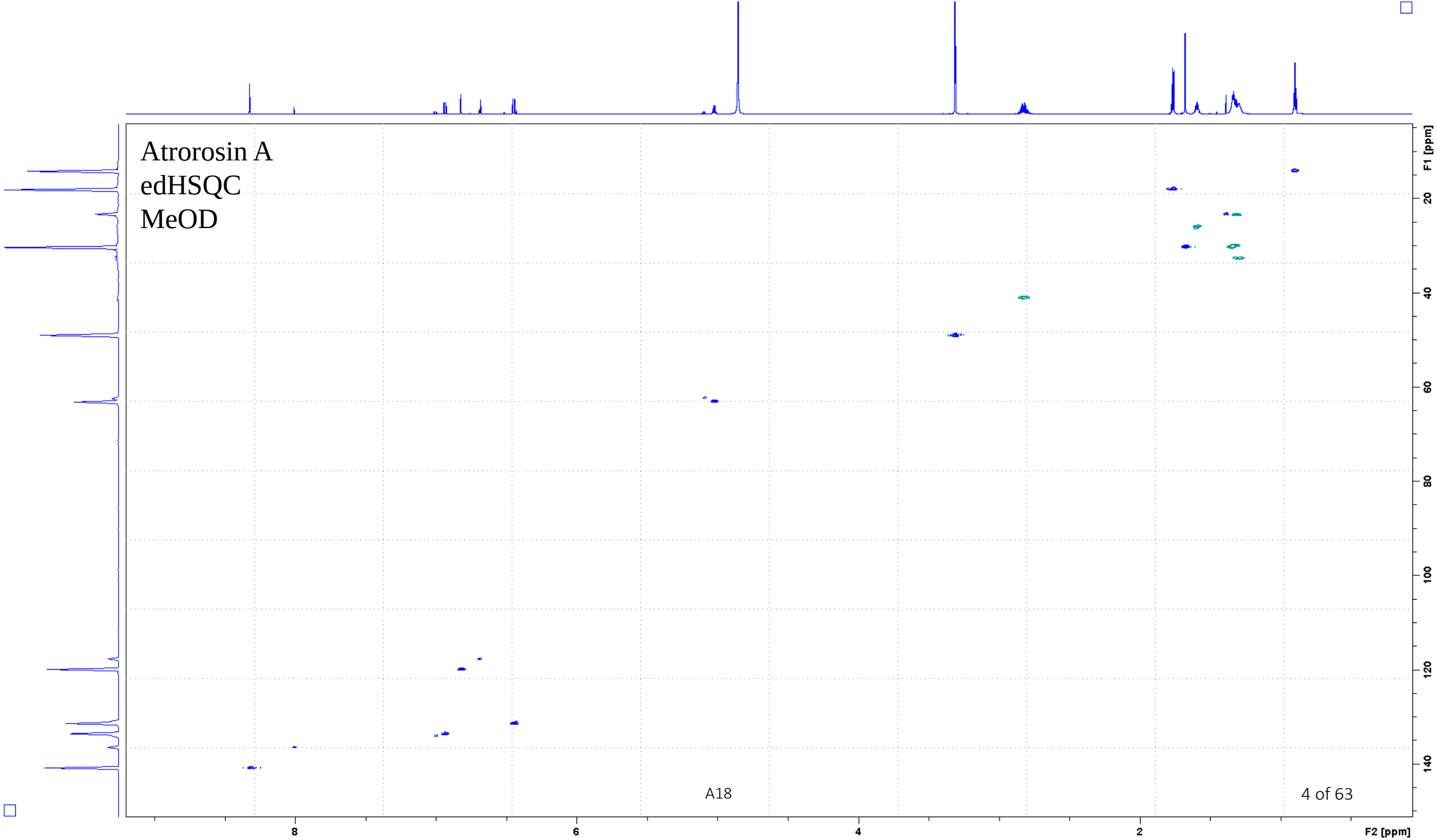
Atrorosin A

^{1}H

MeOD



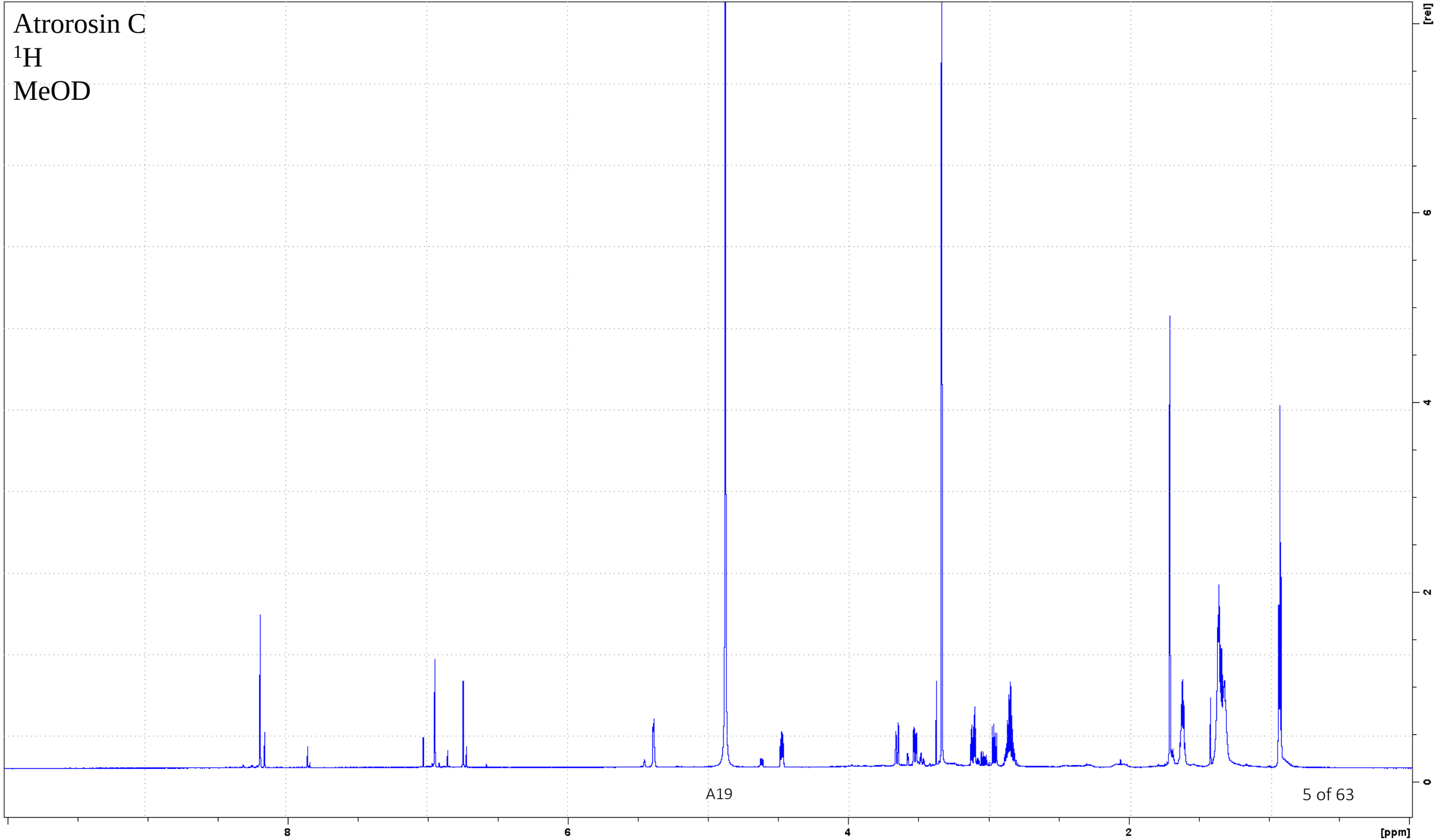
A17

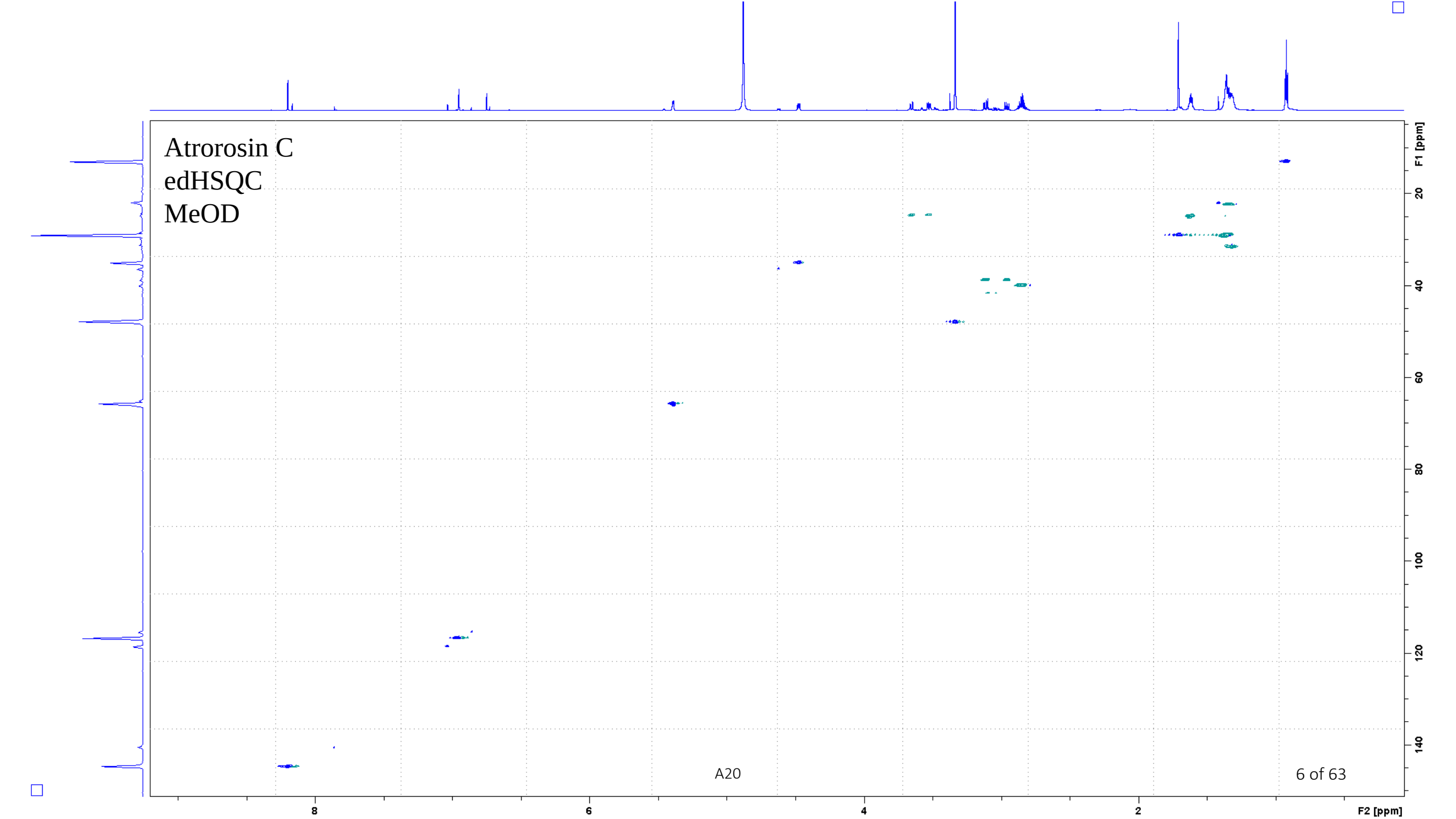

Atrorosin A
edHSQC
MeOD
A18
4 of 63
F1 [ppm]
20
40
60
80
100
120
140
F2 [ppm]
8
6
4
2

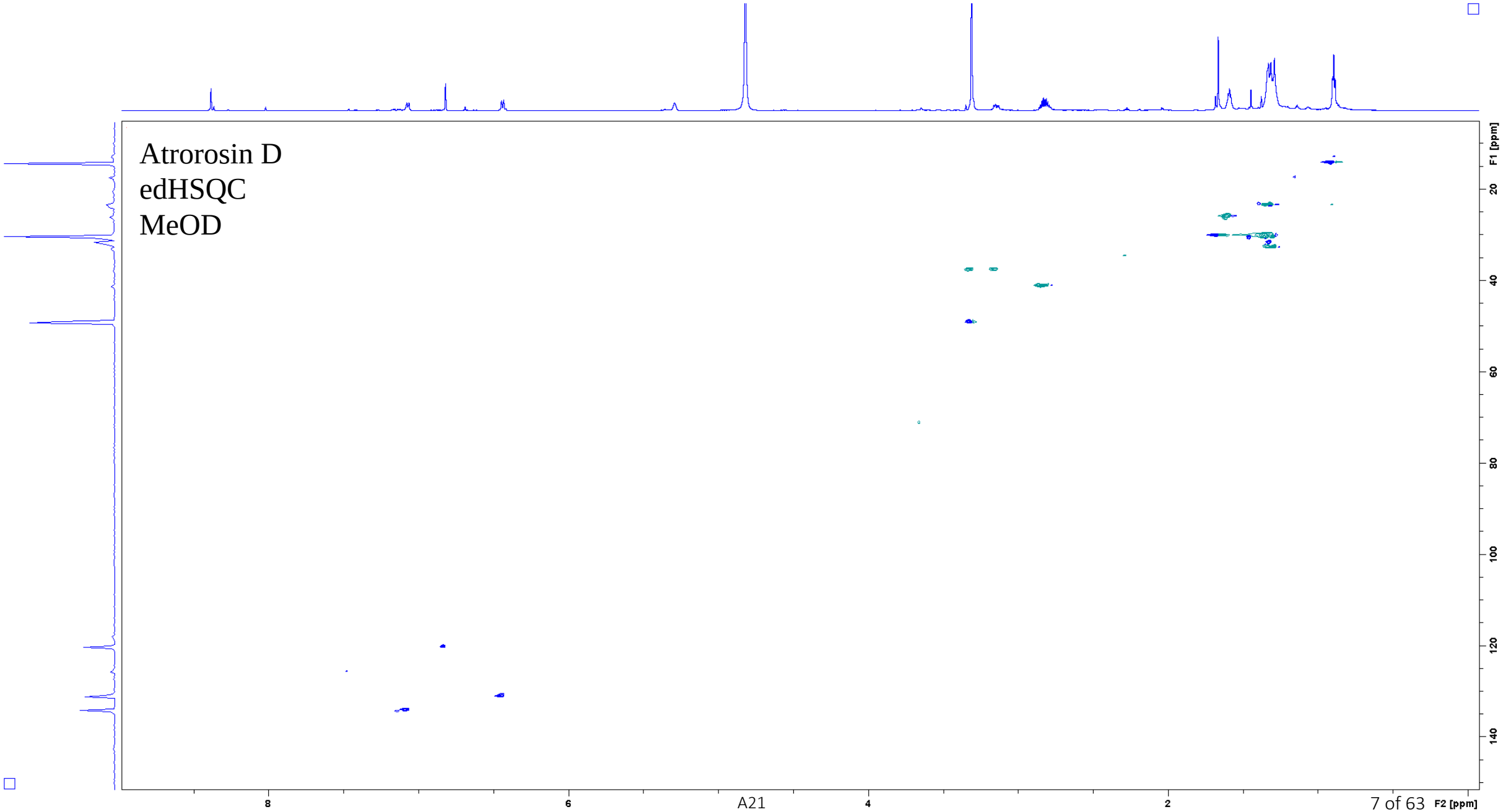
Atrorosin C

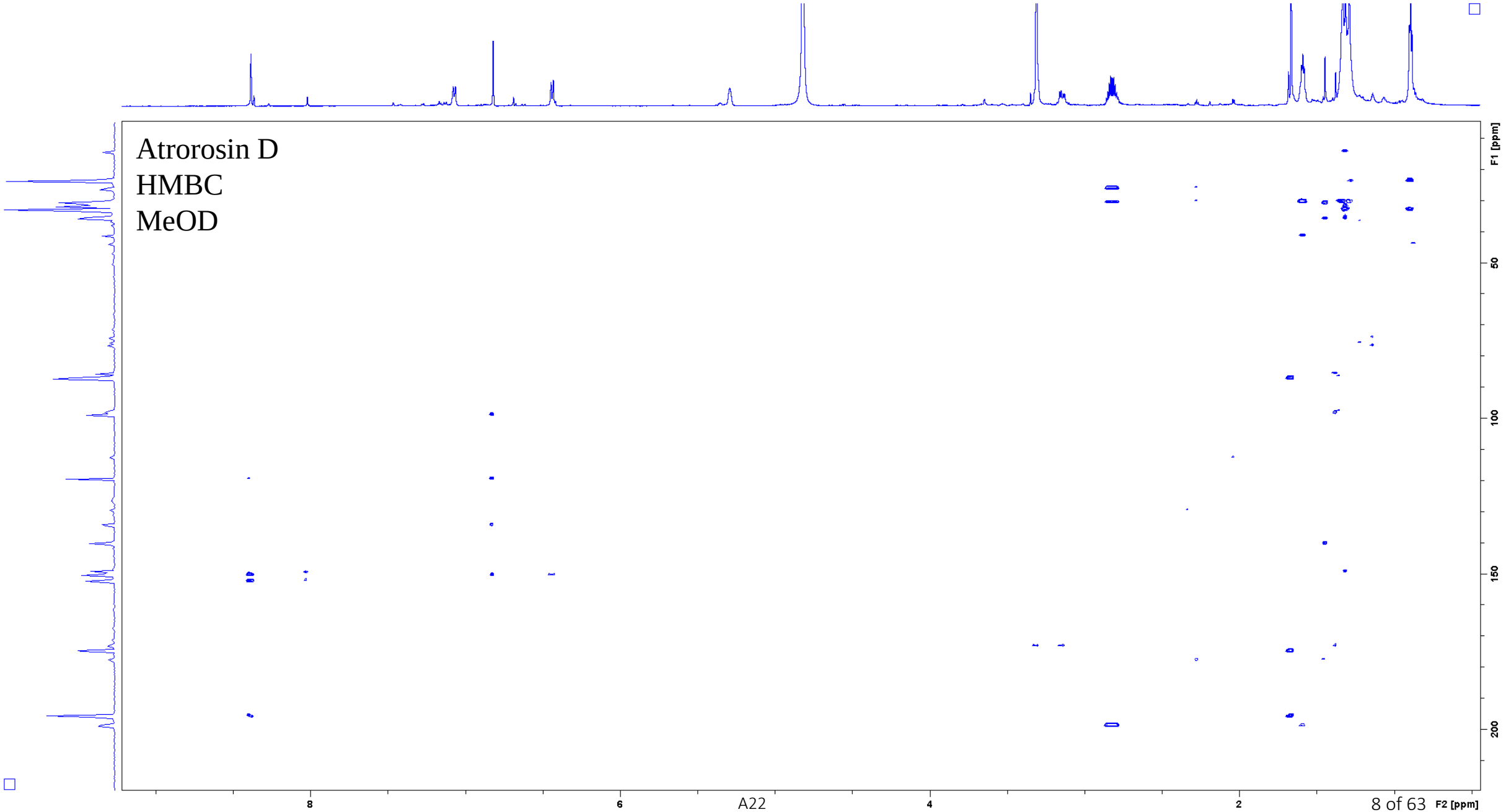
^{1}H

MeOD




Atrorosin C
edHSQC
MeOD
A20
6 of 63
F1 [ppm]
F2 [ppm]
20
40
60
80
100
120
140
8
6
4
2

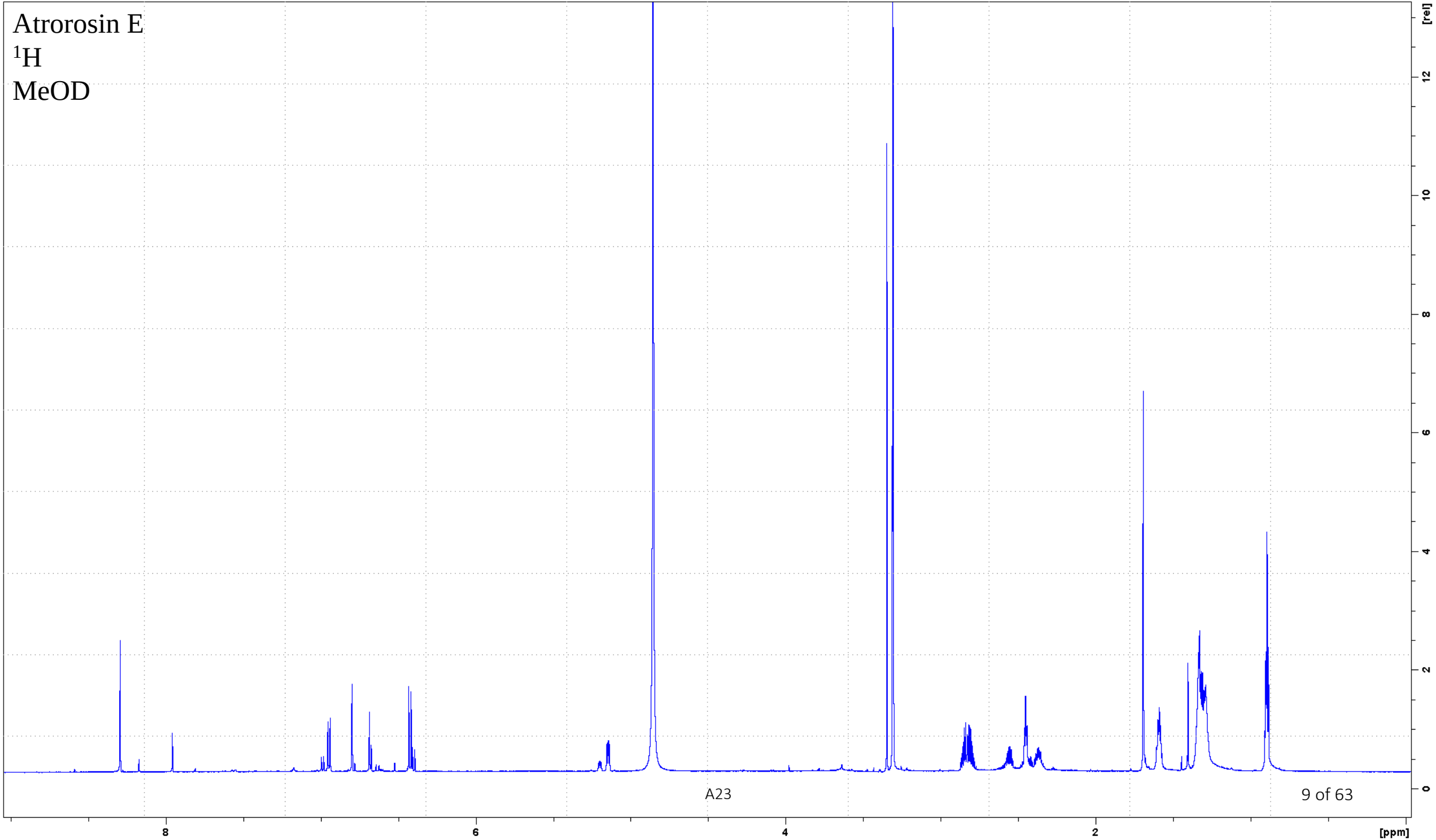

Atrorosin D
edHSQC
MeOD
F1 [ppm]
20
40
60
80
100
120
140
8
6
4
2
F2 [ppm]


Atrorosin D
HMBC
MeOD
F1 [ppm]
50
100
150
200
8
6
4
2
F2 [ppm]

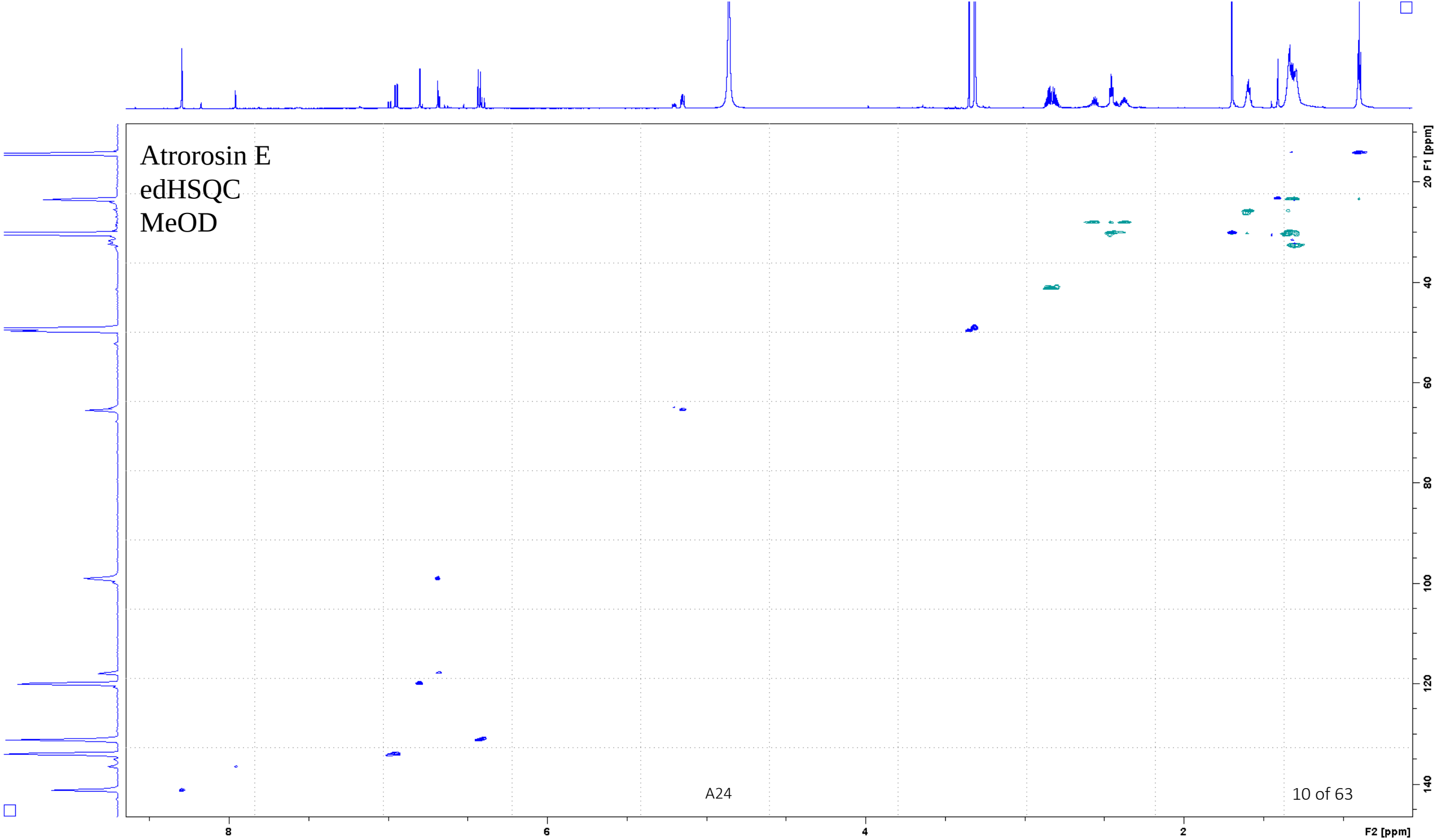
Atrorosin E

^{1}H

MeOD



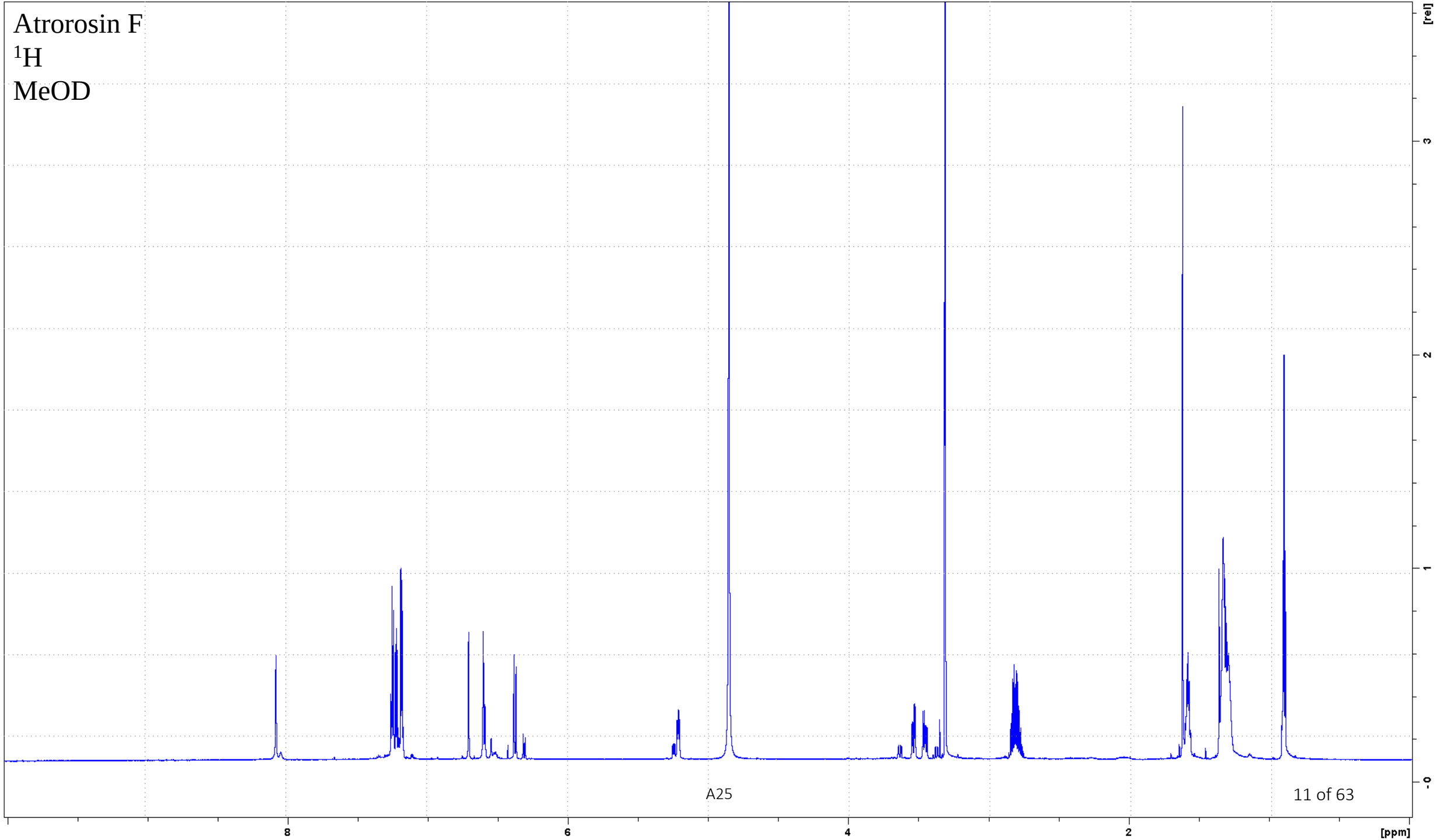
A23


Atrorosin E
edHSQC
MeOD
A24
10 of 63
F2 [ppm]
F1 [ppm]

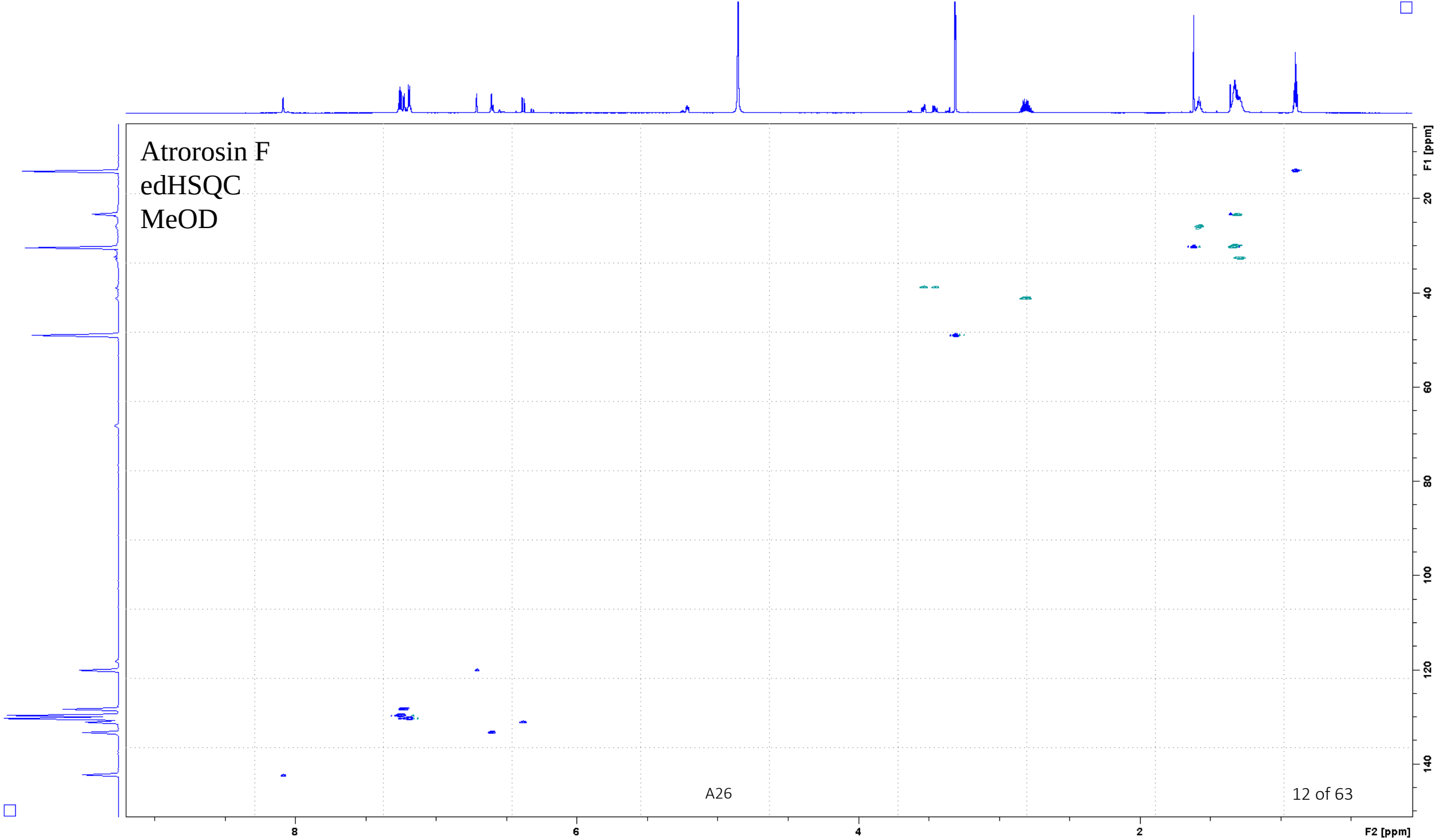
Atrorosin F

^{1}H

MeOD



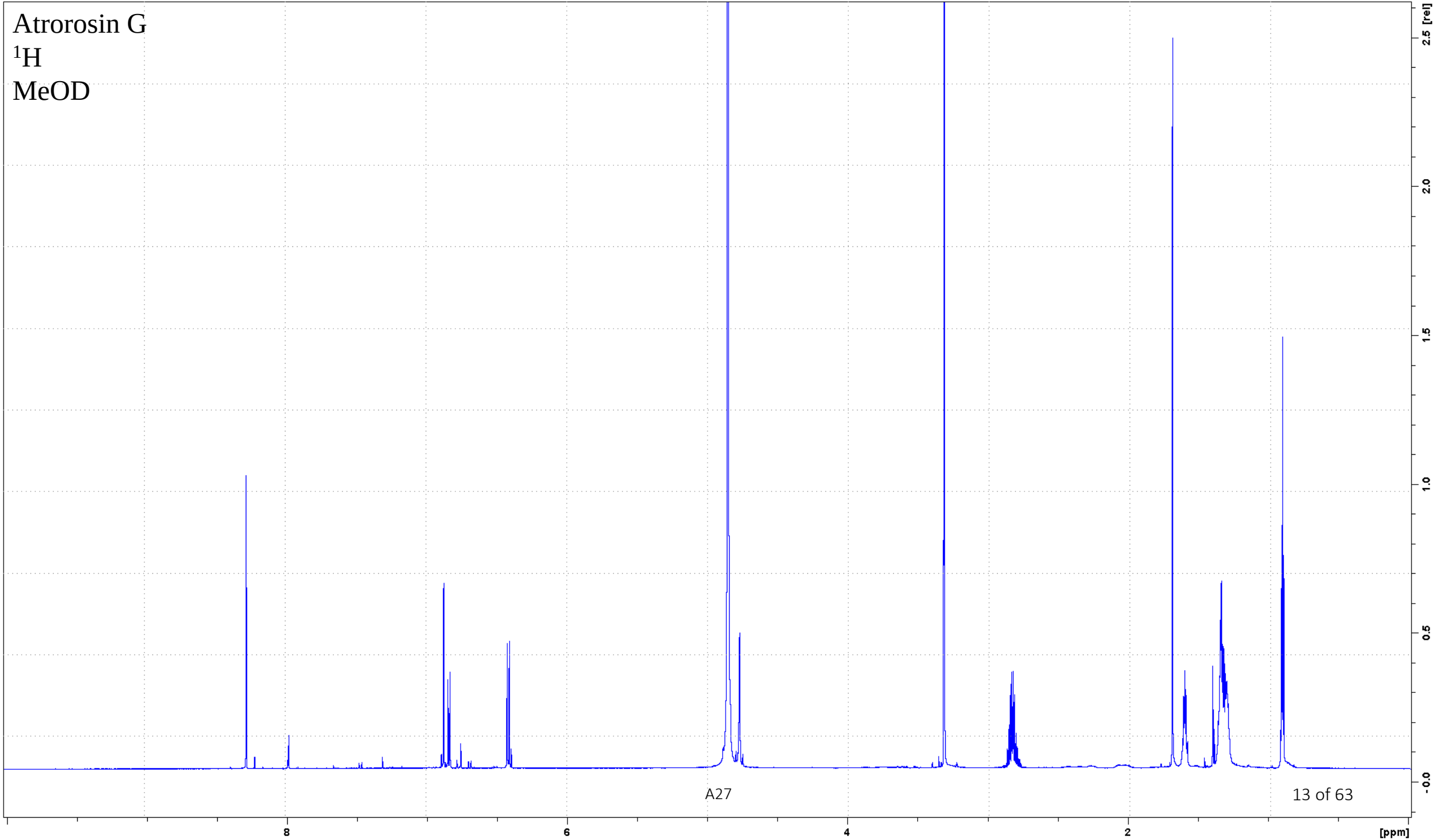
A25


Atrorosin F
edHSQC
MeOD
A26
12 of 63
F1 [ppm]
F2 [ppm]
20
40
60
80
100
120
140
8
6
4
2

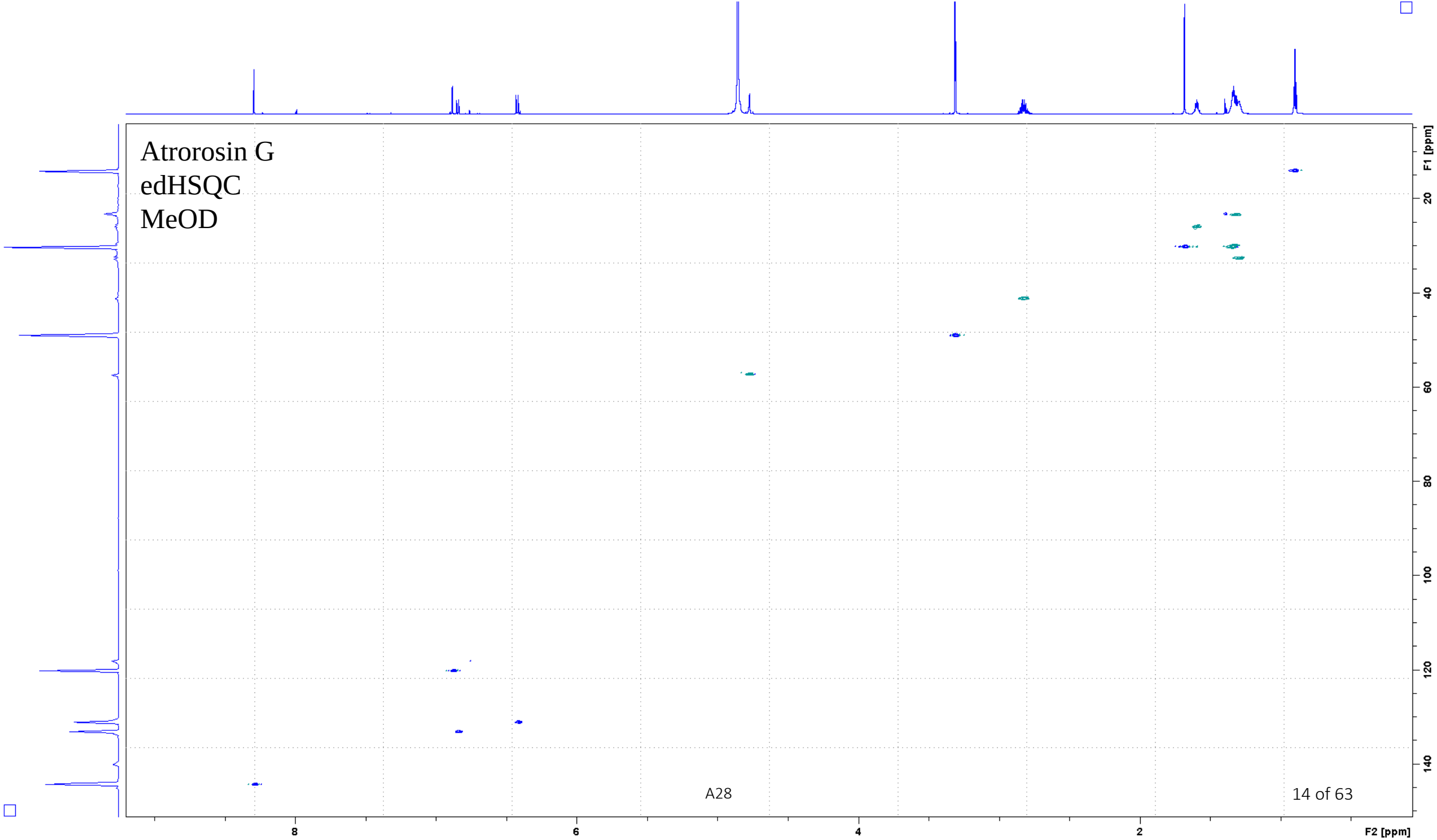
Atrorosin G

^{1}H

MeOD



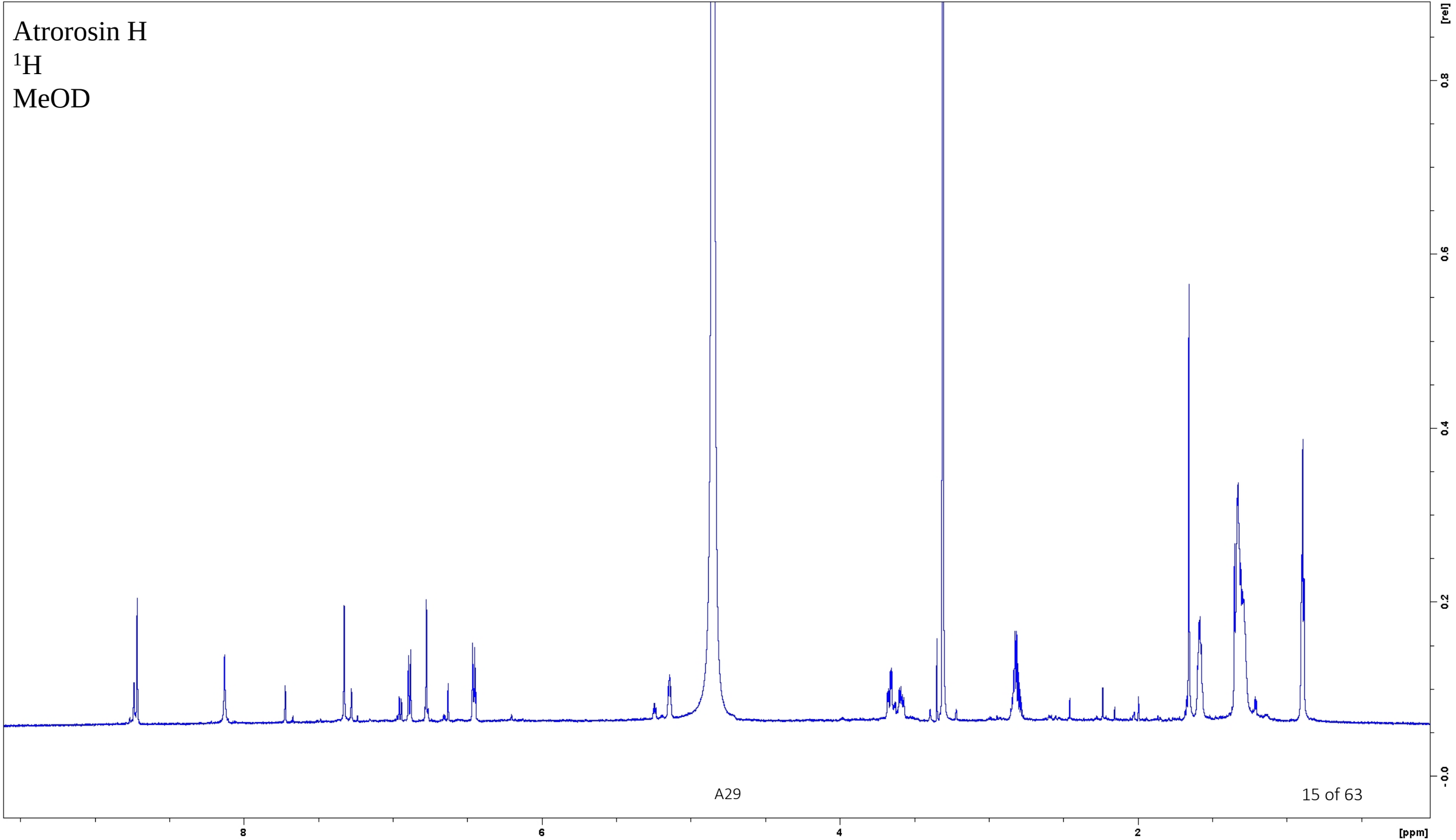
A27

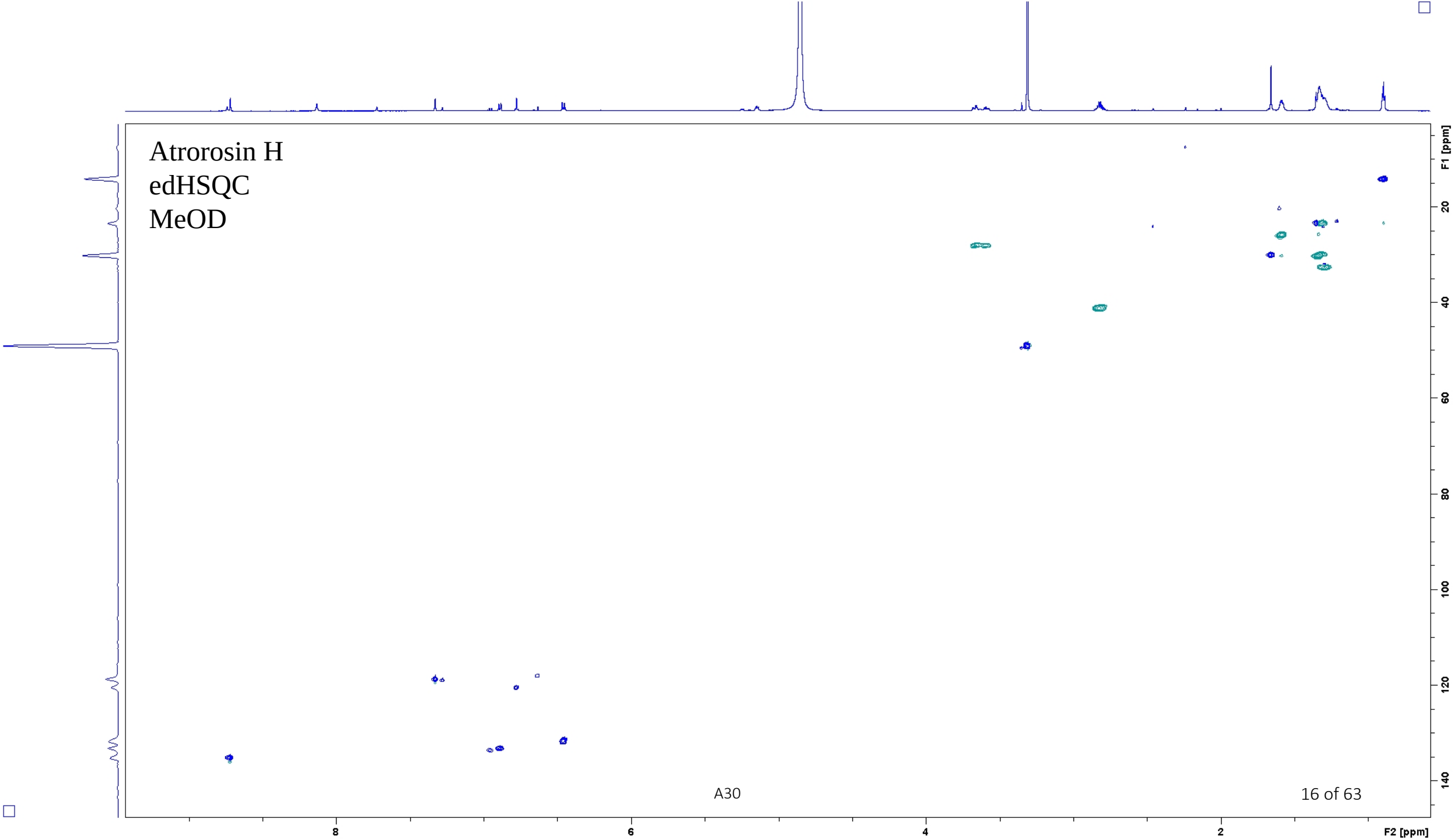

Atrorosin G
edHSQC
MeOD
A28
14 of 63
F1 [ppm]
F2 [ppm]
20
40
60
80
100
120
140
8
6
4
2

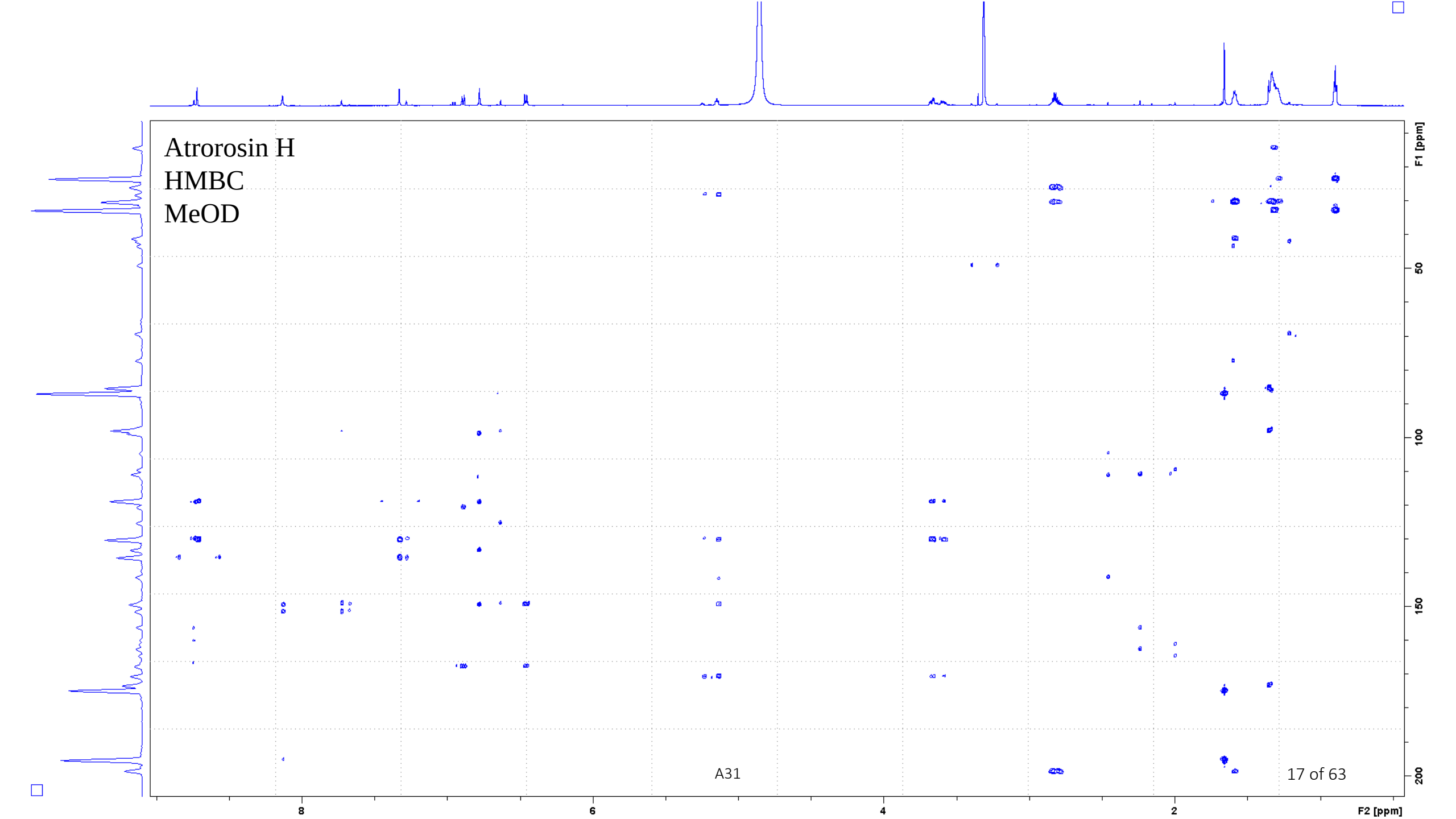
Atrorosin H

^{1}H

MeOD



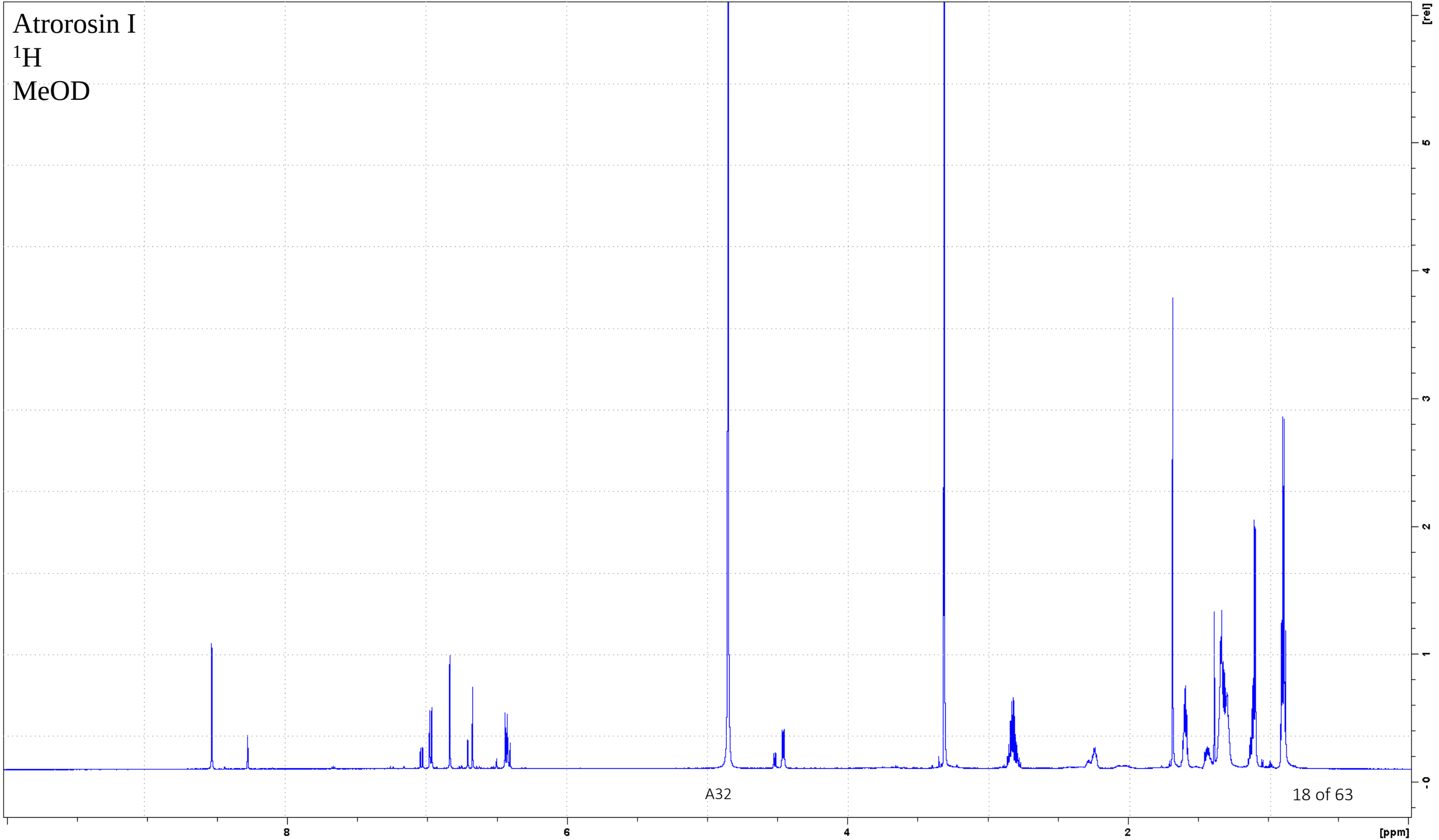

Atrorosin H
edHSQC
MeOD
F1 [ppm]
20
40
60
80
100
120
140
8
6
4
2
F2 [ppm]


Atrorosin H
HMBC
MeOD
A31
17 of 63
F1 [ppm]
F2 [ppm]
50
100
150
200
8
6
4
2

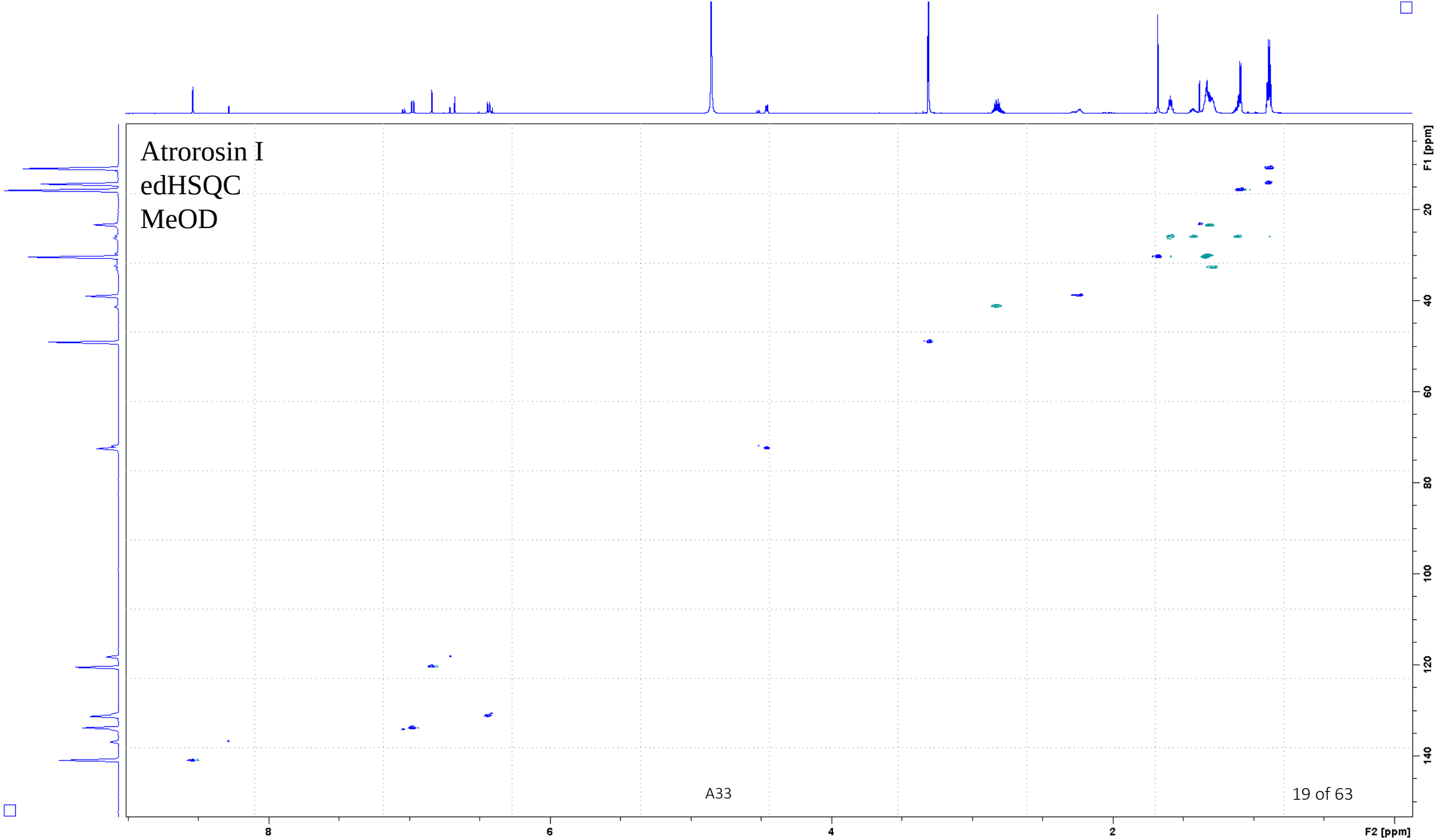
Atrorosin I

^{1}H

MeOD



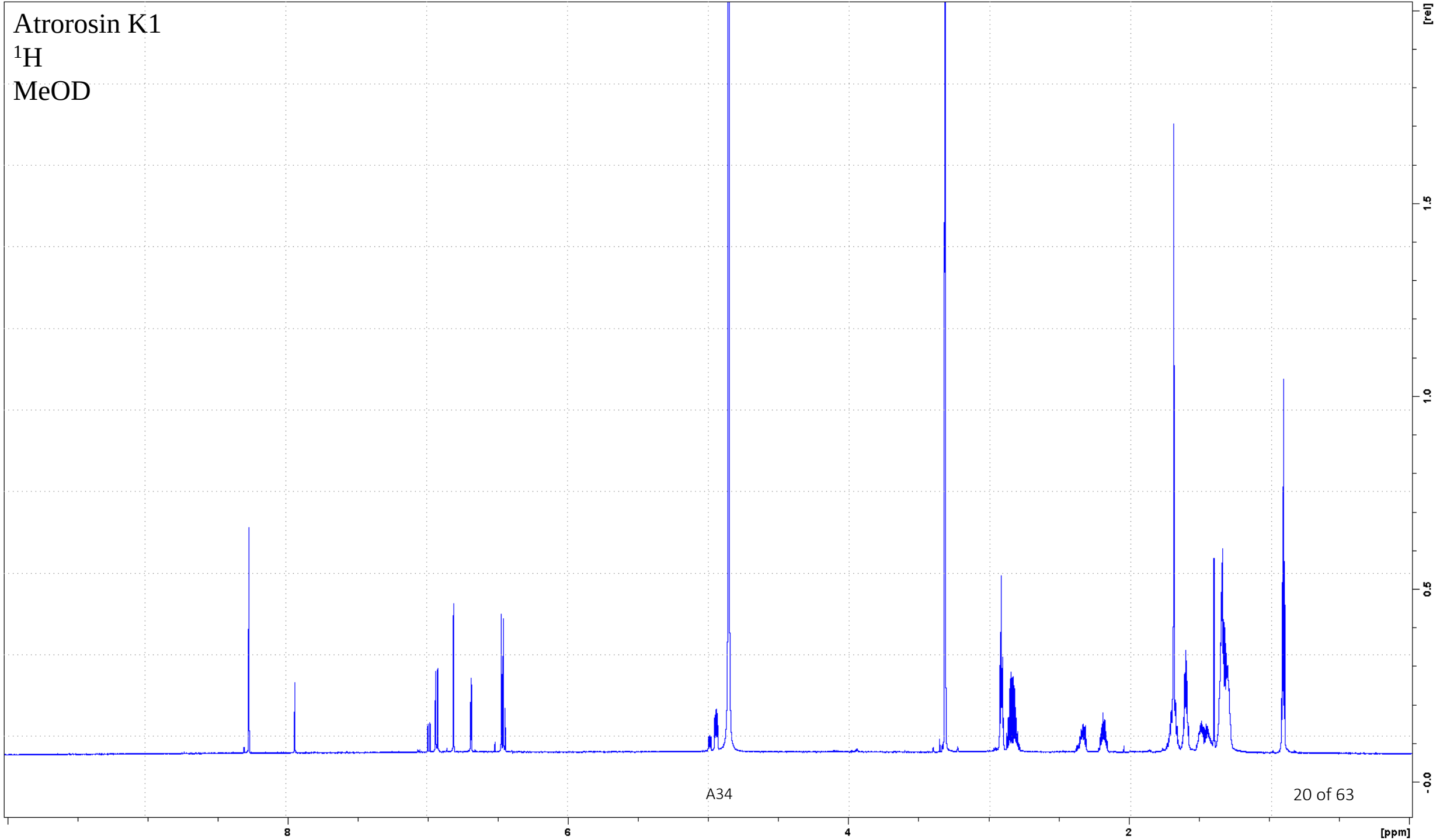
A32


Atrorosin I
edHSQC
MeOD
F1 [ppm]
20
40
60
80
100
120
140
8
6
4
2
F2 [ppm]

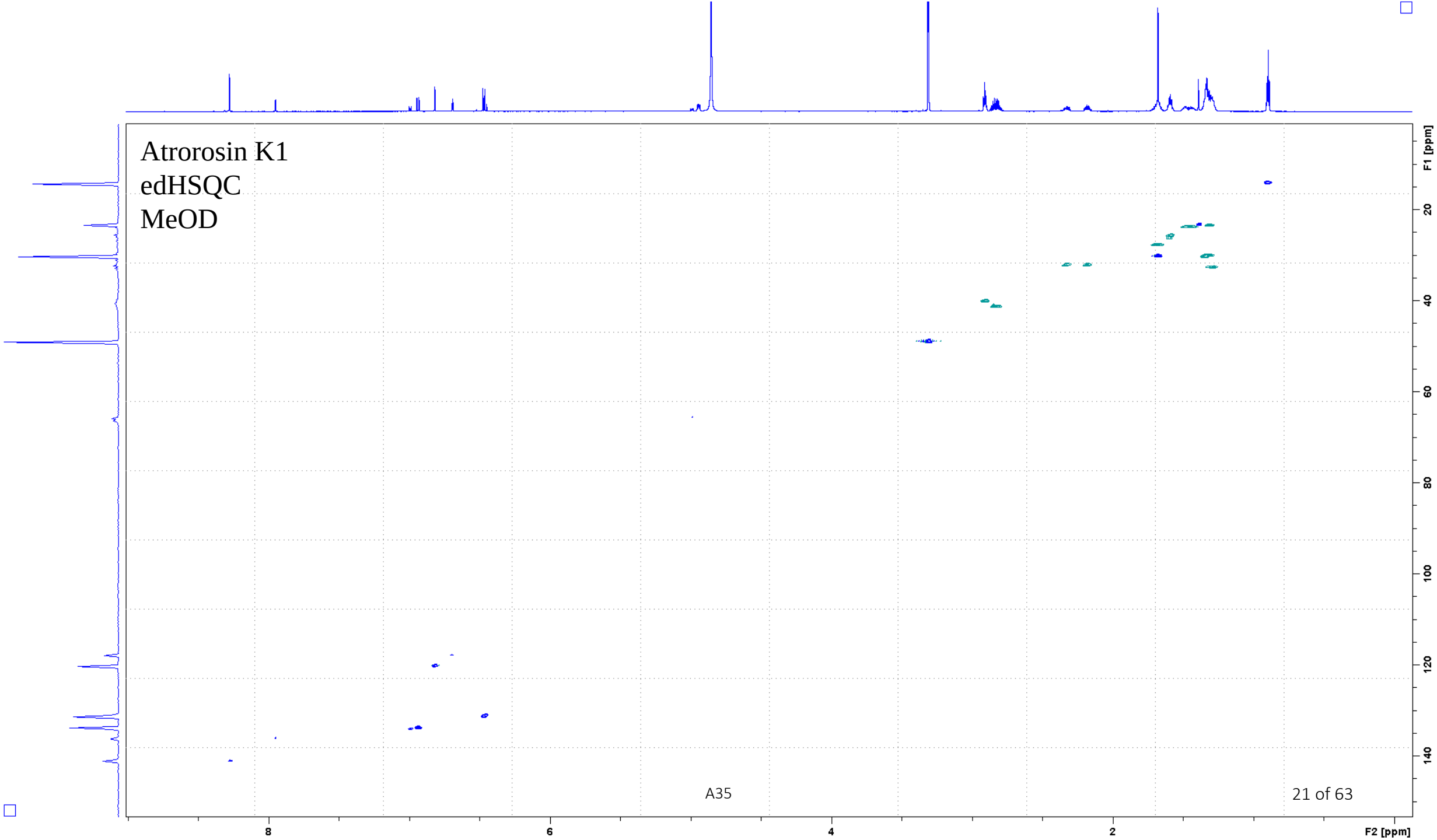
Atrorosin K1

^{1}H

MeOD



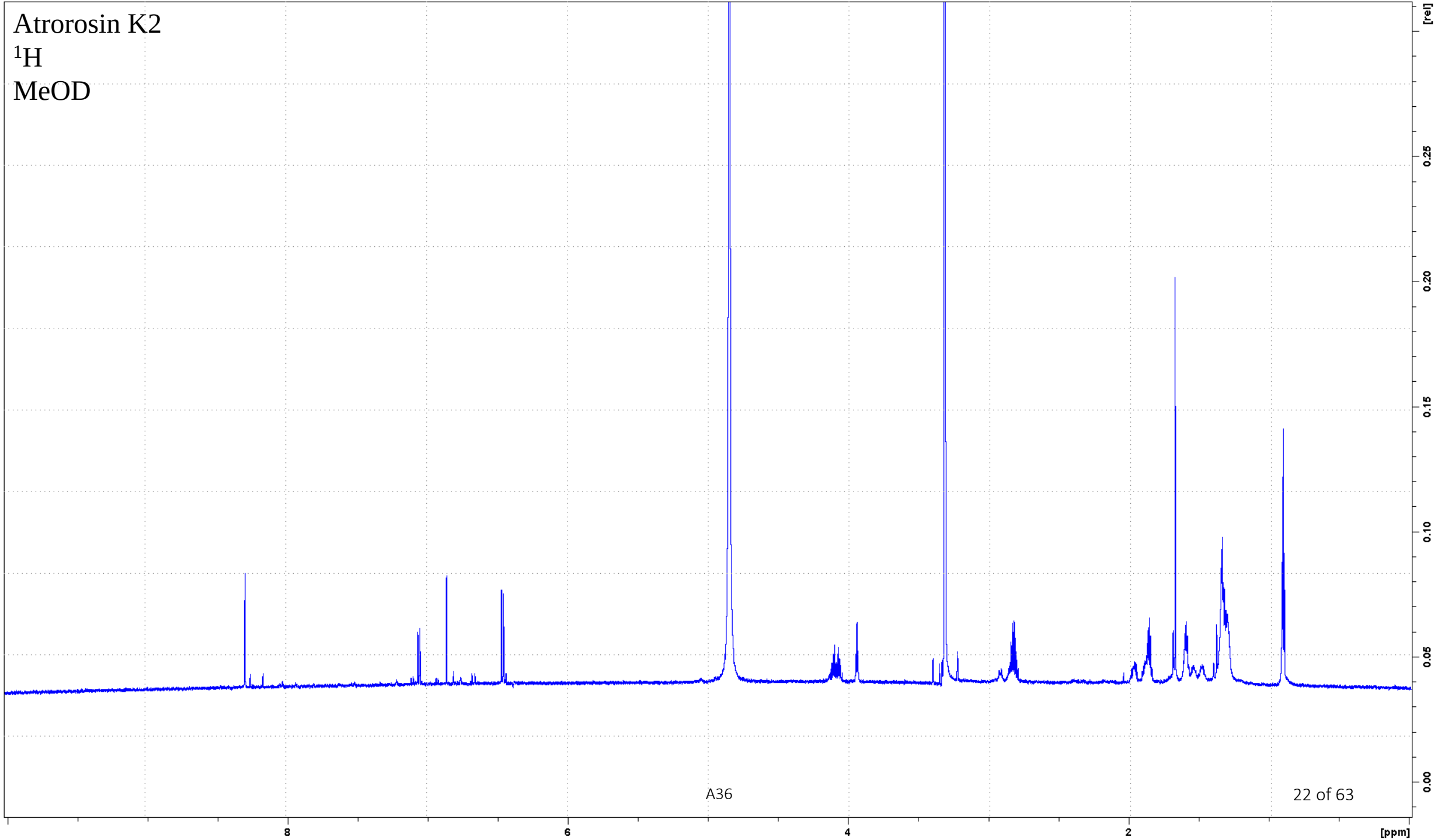
A34


Atrorosin K1
edHSQC
MeOD
A35
21 of 63
F1 [ppm]
F2 [ppm]
20
40
60
80
100
120
140
8
6
4
2

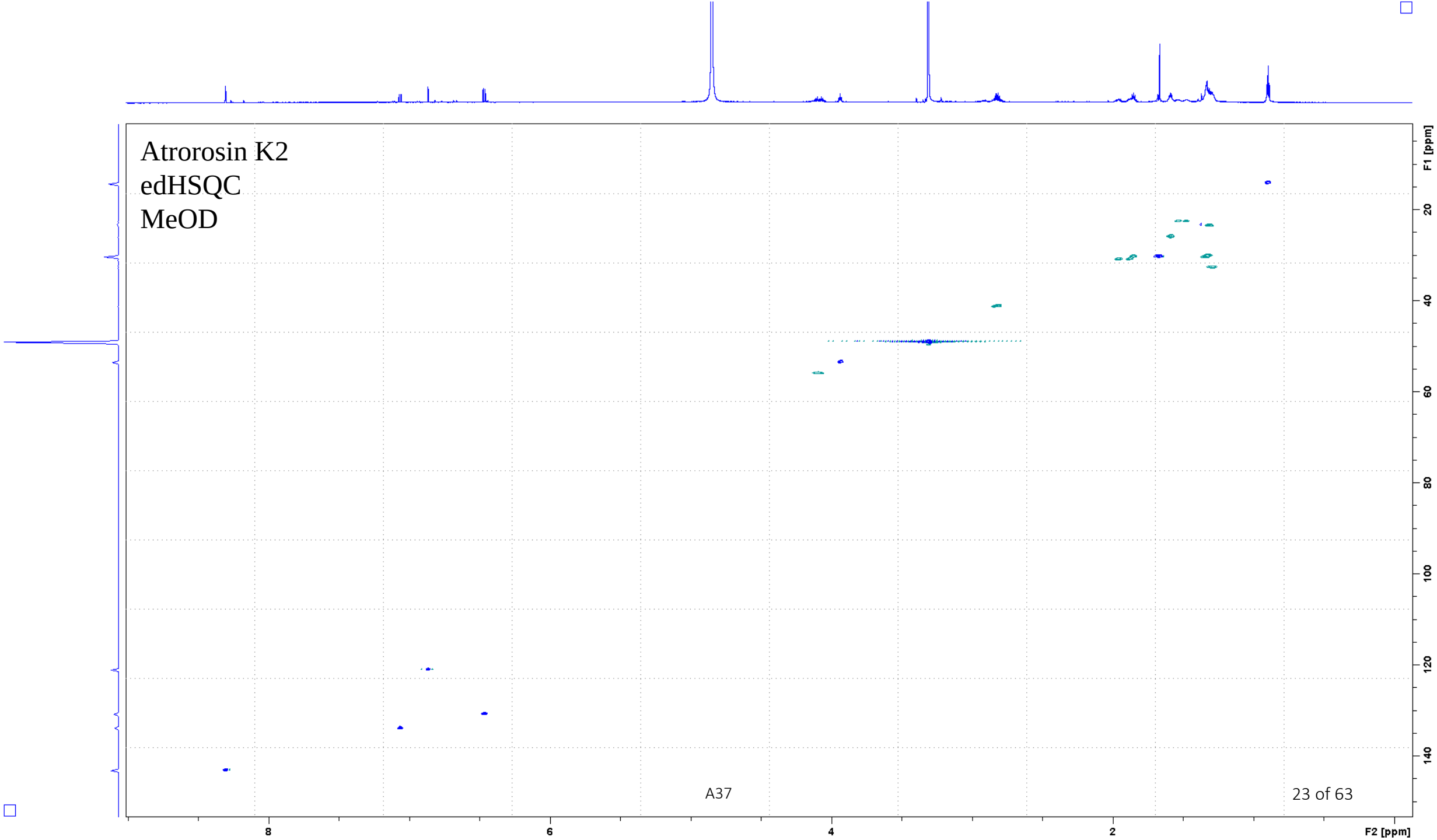
Atrorosin K2

^{1}H

MeOD



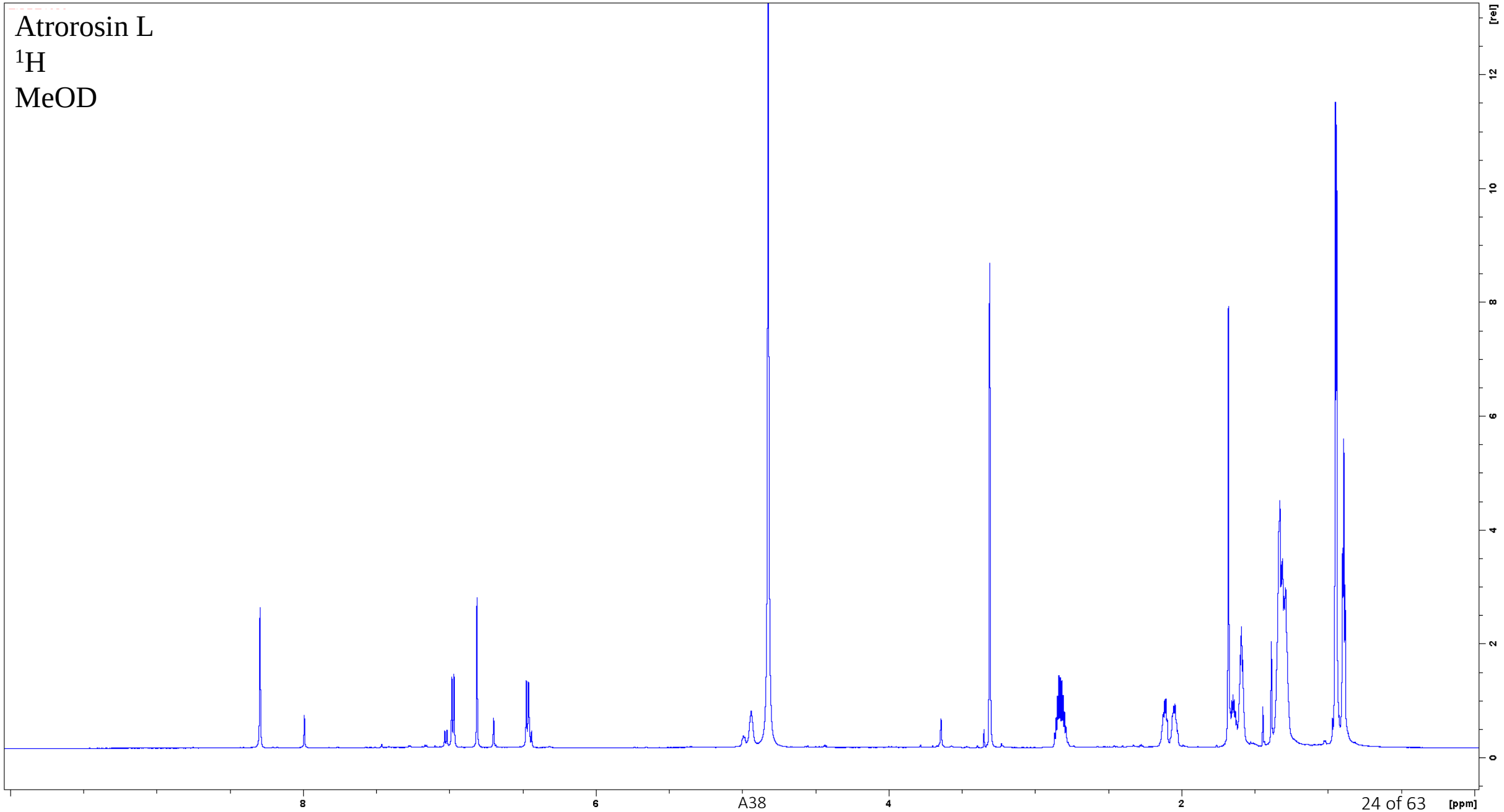
A36

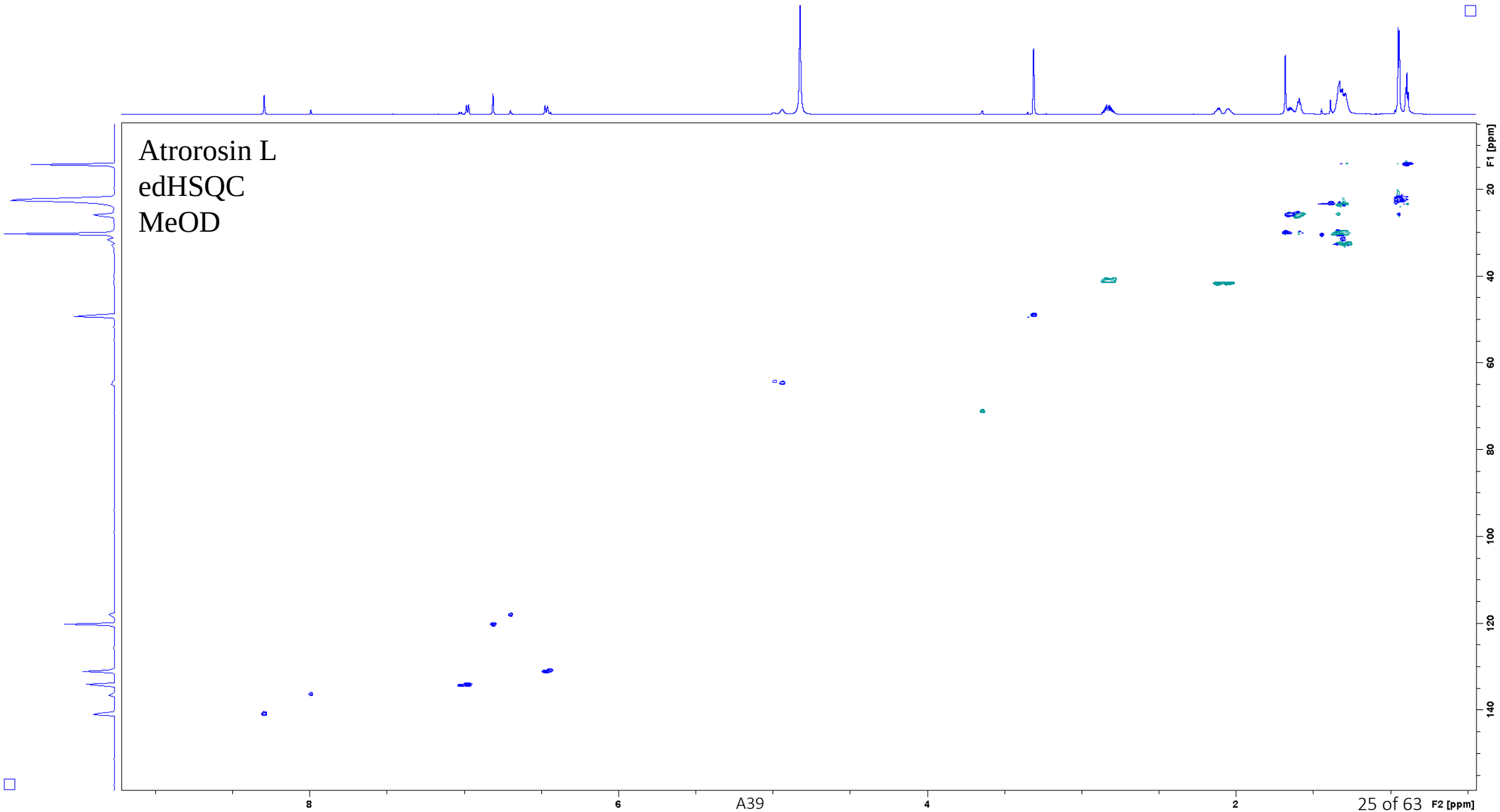

Atrorosin K2
edHSQC
MeOD
A37
23 of 63
F1 [ppm]
F2 [ppm]
20
40
60
80
100
120
140
8
6
4
2

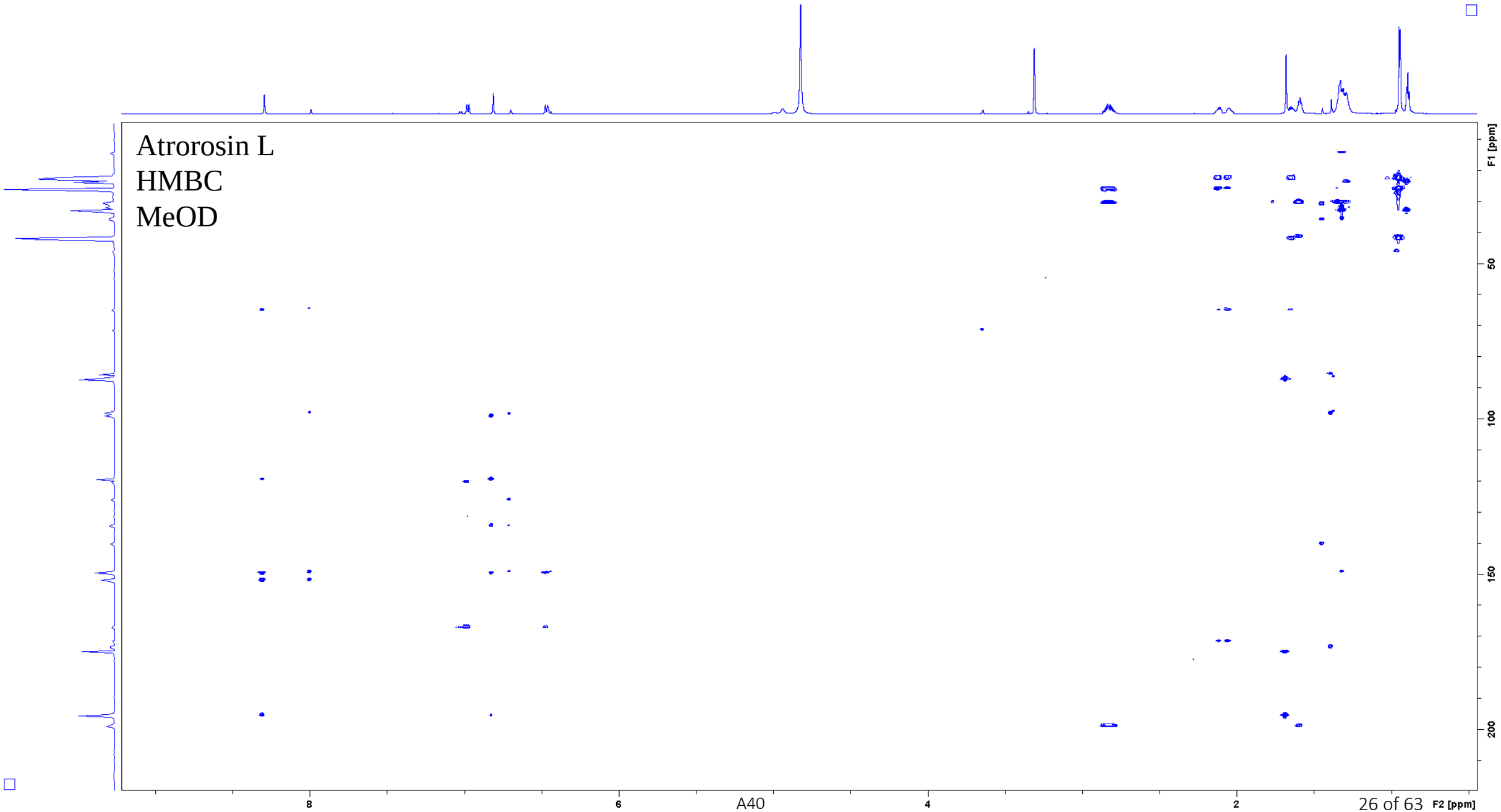
Atrorosin L

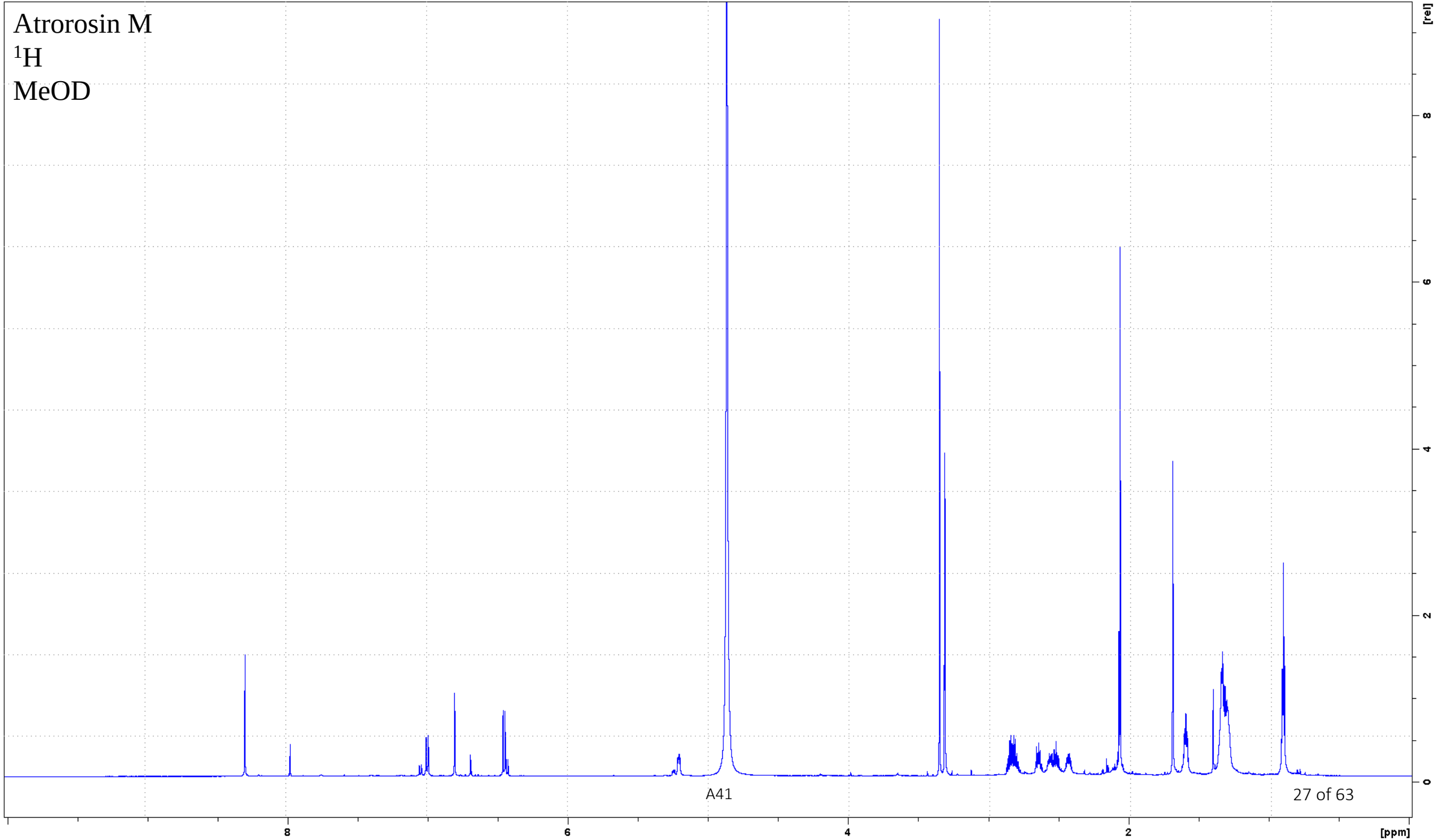
^{1}H

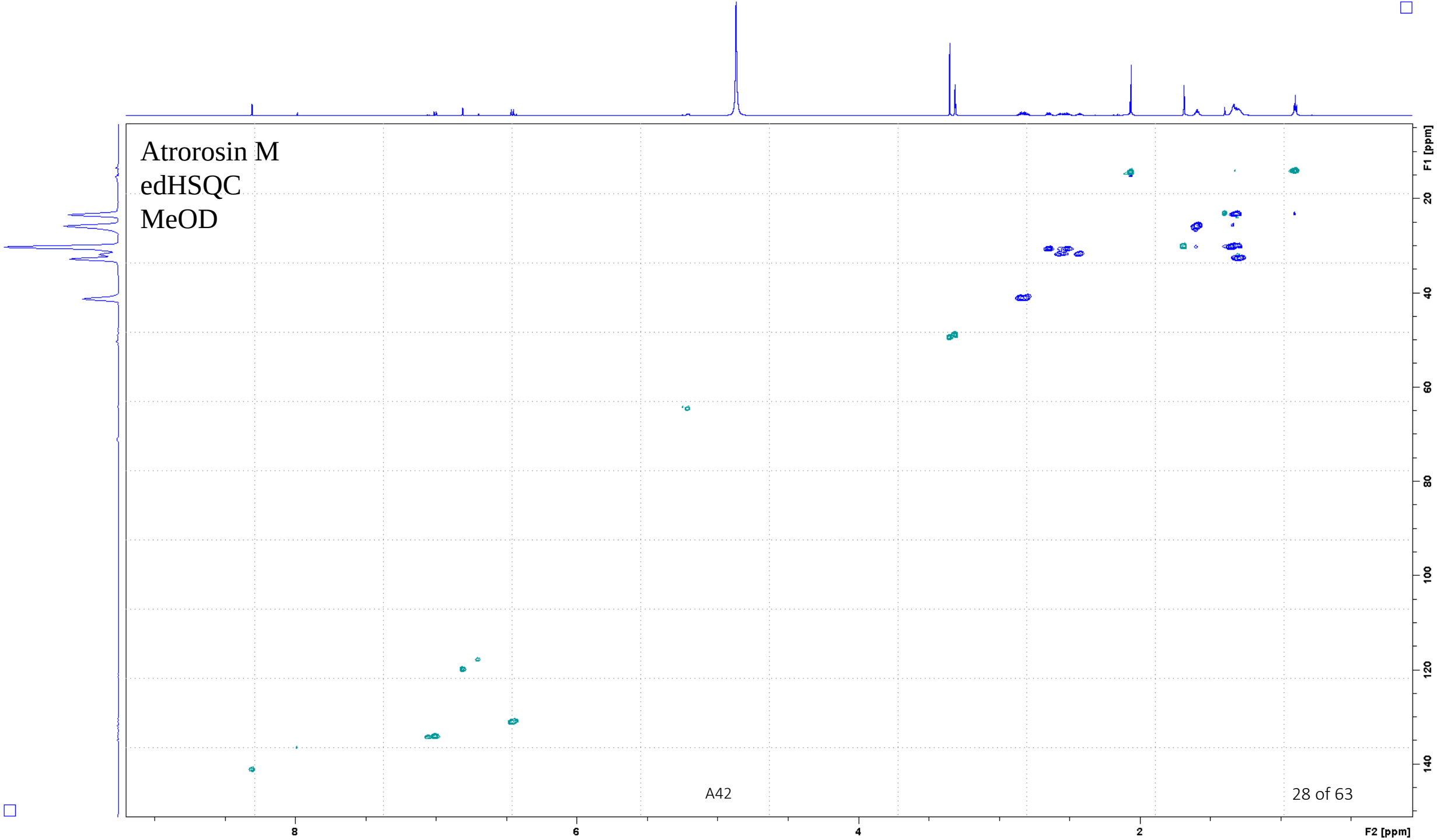
MeOD

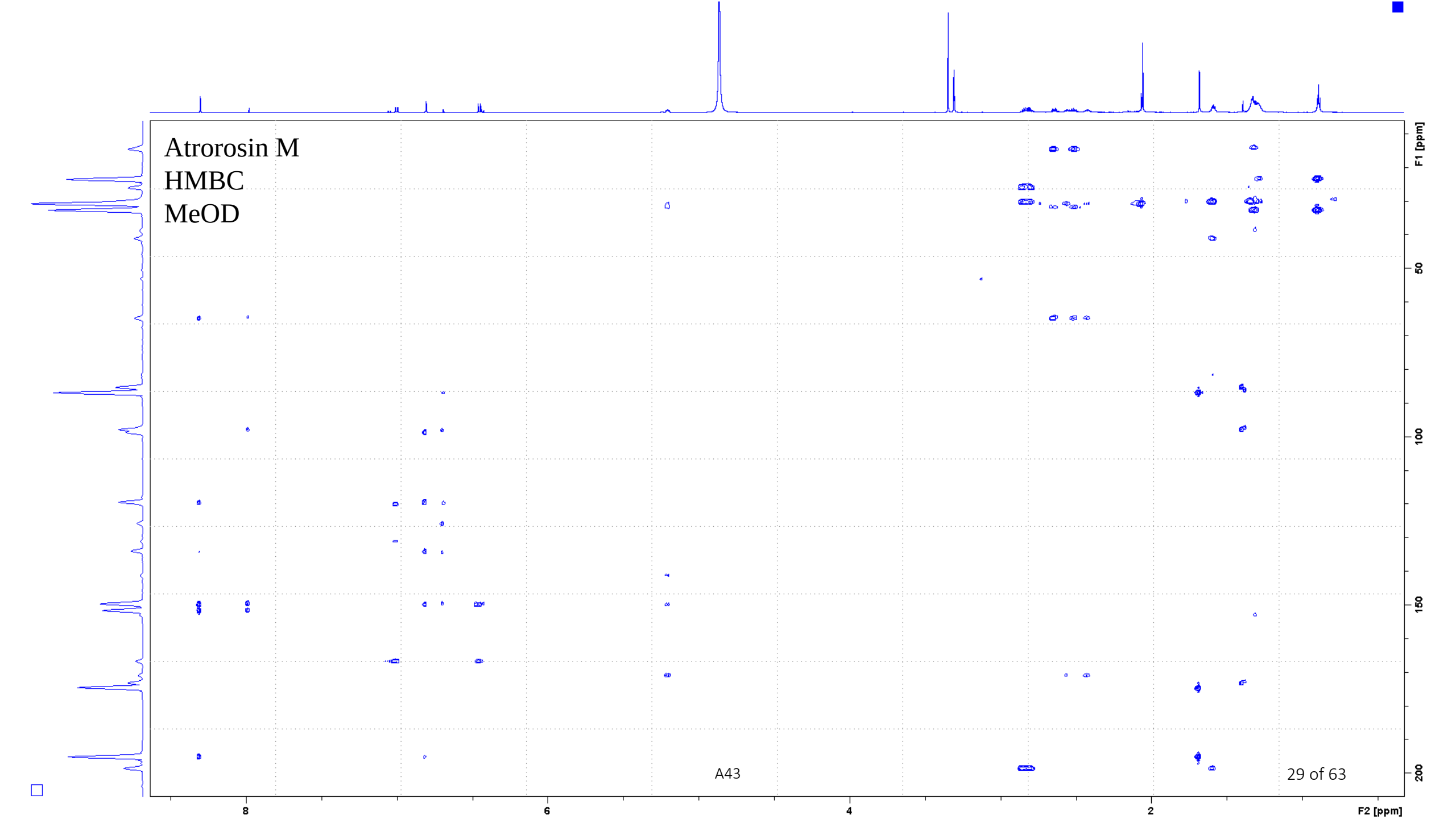



Atrorosin L
edHSQC
MeOD
F1 [ppm]
20
40
60
80
100
120
140
8
6
4
2
F2 [ppm]


Atrorosin L
HMBC
MeOD
F1 [ppm]
50
100
150
200
8
6
4
2
F2 [ppm]


Atrorosin M
1H
MeOD
A41
27 of 63
8
6
4
2
[ppm]
[rel]
8
6
4
2
0

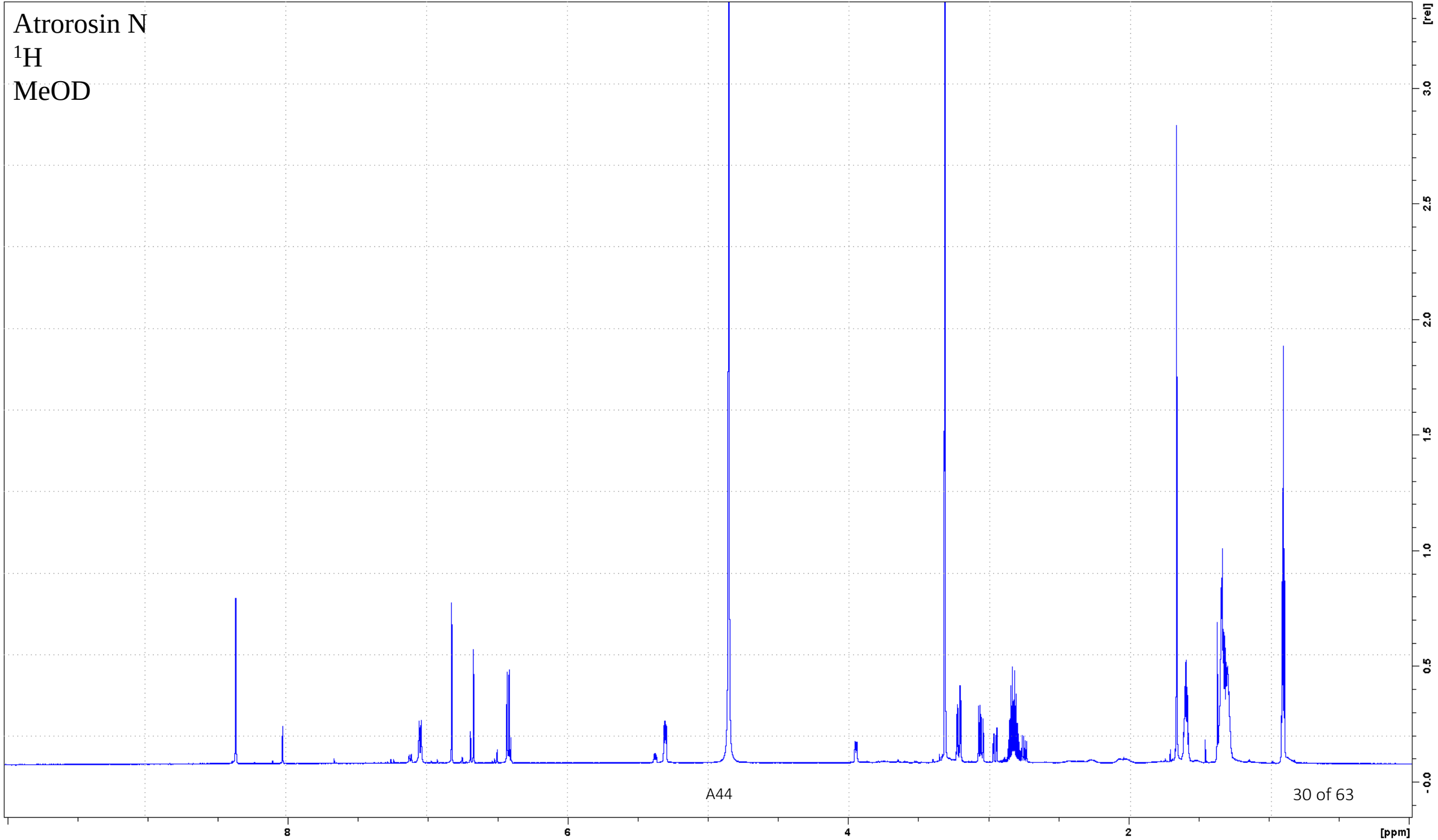

Atrorosin M
edHSQC
MeOD
F1 [ppm]
20
40
60
80
100
120
140
8
6
4
2
F2 [ppm]


Atrorosin M
HMBC
MeOD
A43
29 of 63
F1 [ppm]
F2 [ppm]
8
6
4
2
50
100
150
200

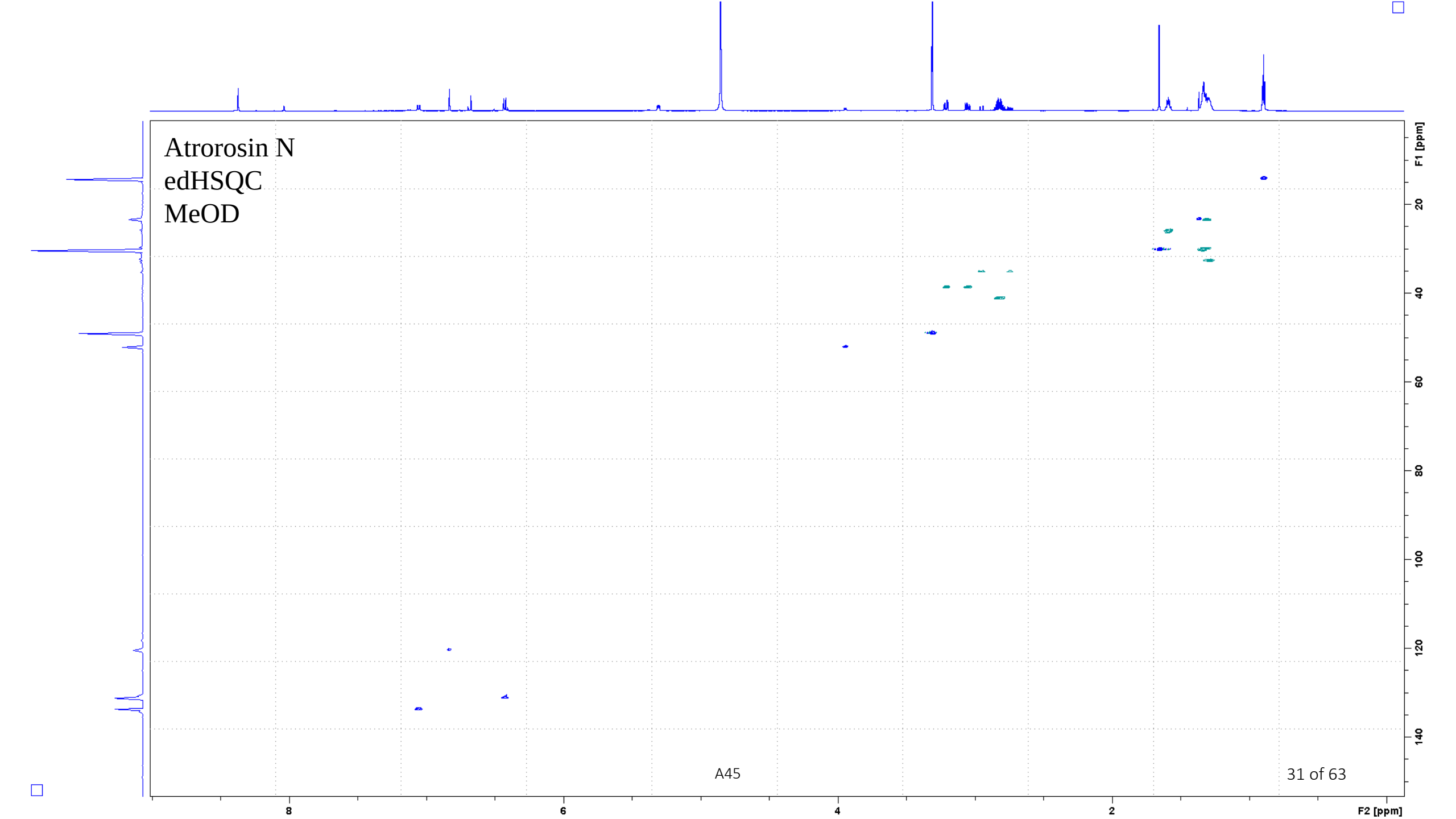
Atrorosin N

^{1}H

MeOD



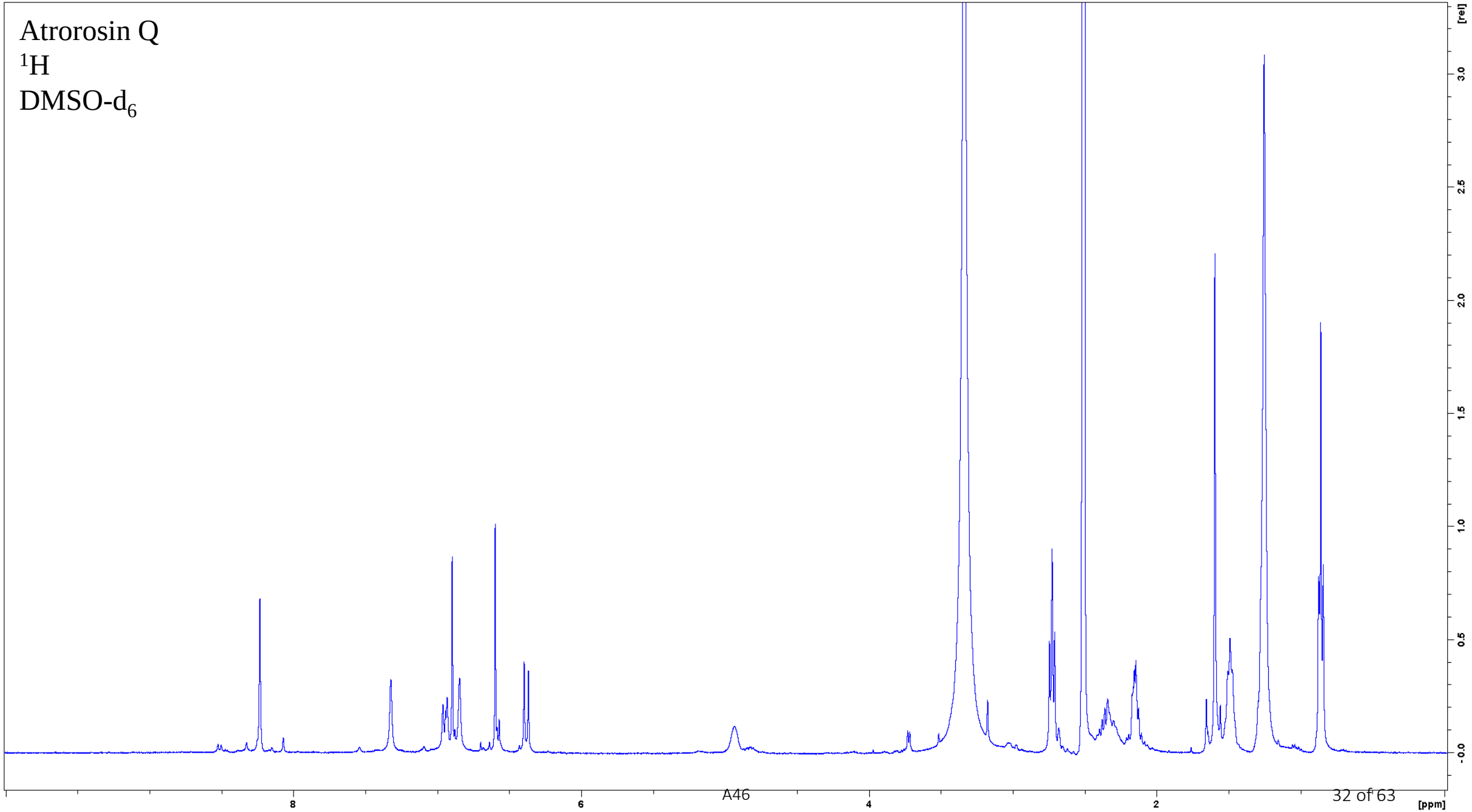
A44

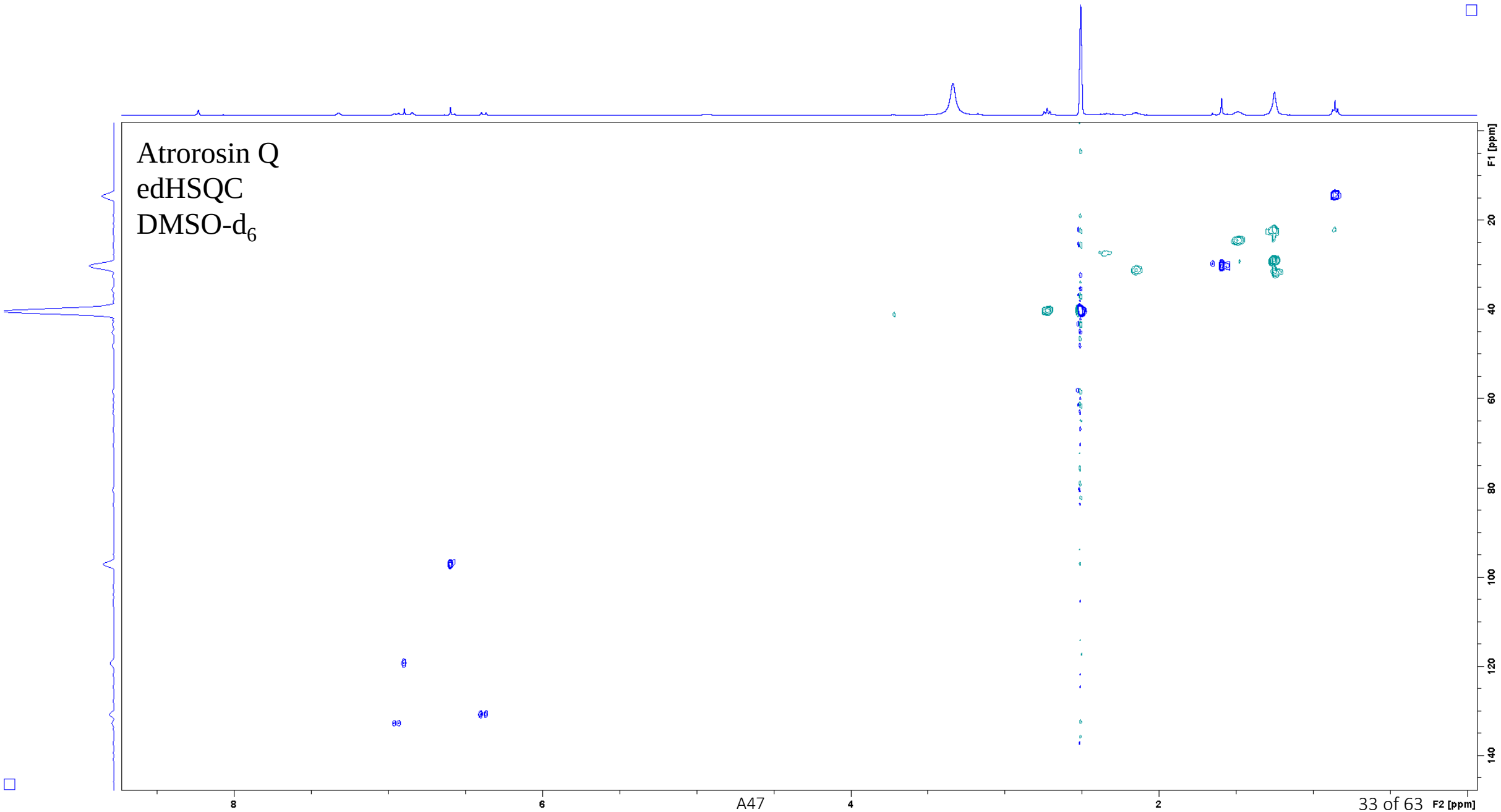

Atrorosin N
edHSQC
MeOD
A45
31 of 63
F1 [ppm]
F2 [ppm]
20
40
60
80
100
120
140
8
6
4
2

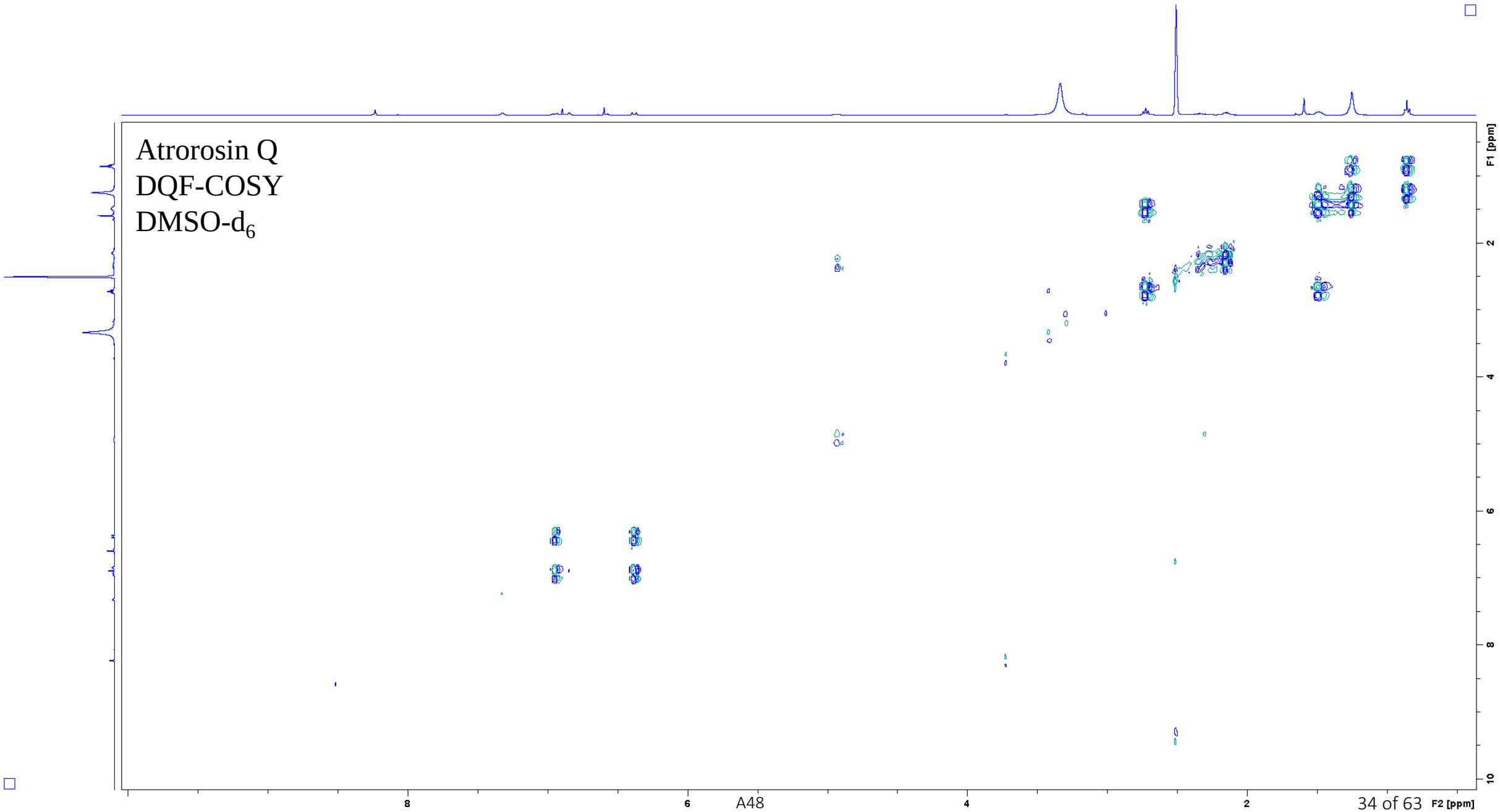
Atrorosin Q

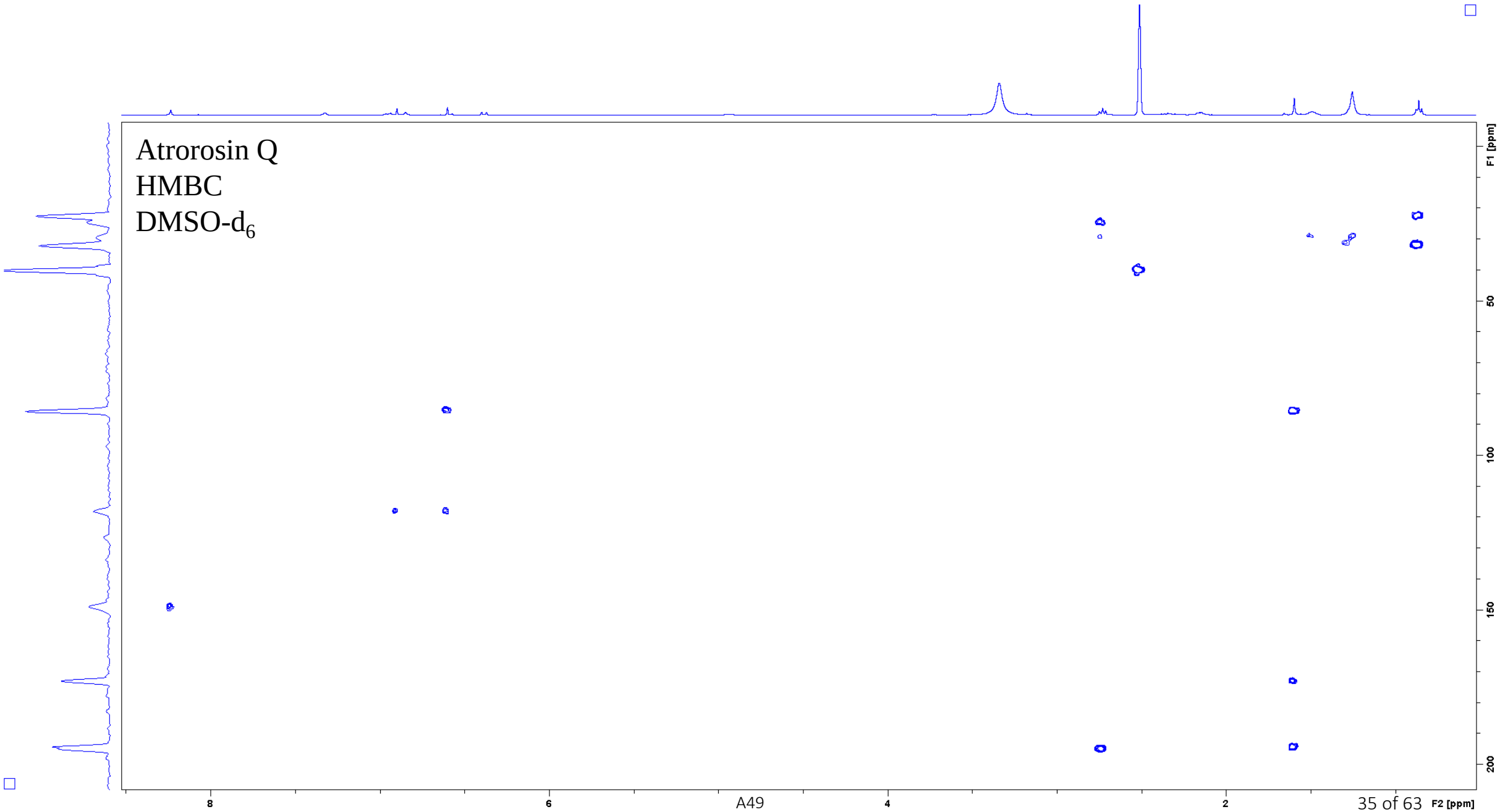
^{1}H

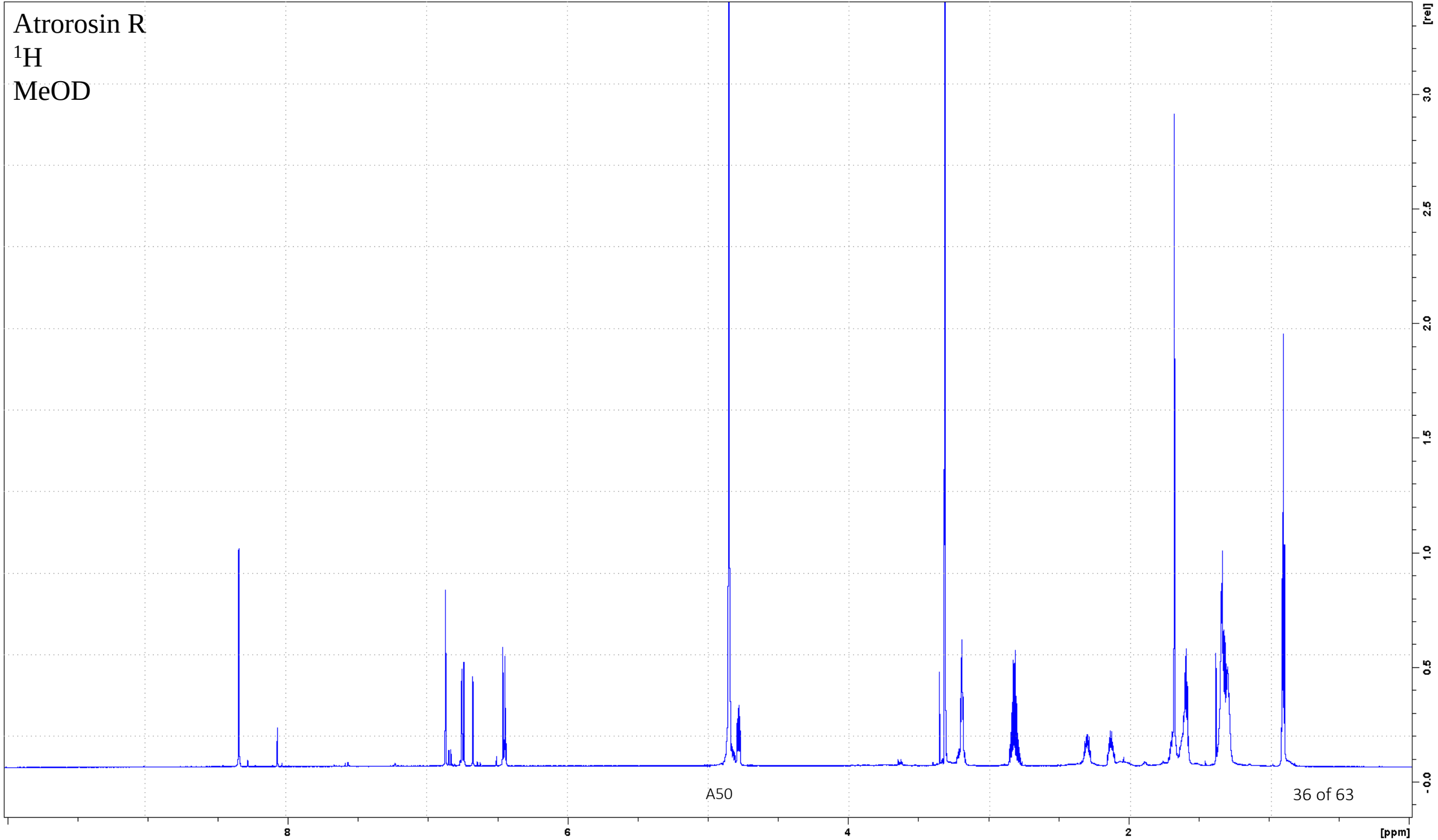
DMSO-d_6

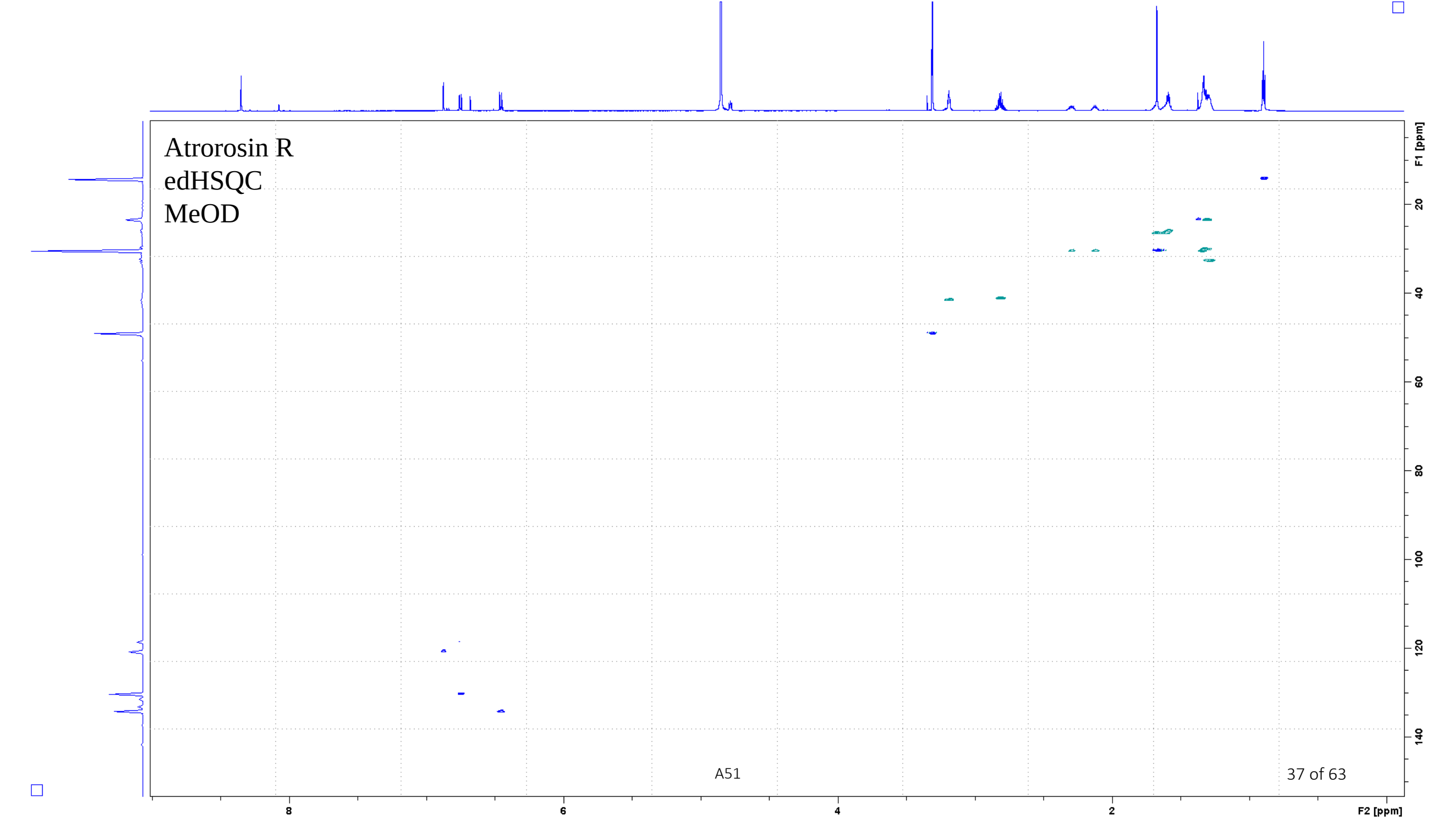



Atrorosin Q
edHSQC
DMSO-d6
F1 [ppm]
20
40
60
80
100
120
140
8
6
4
2
F2 [ppm]


Atrorosin Q
DQF-COSY
DMSO-d6
F1 [ppm]
F2 [ppm]
8
6
4
2
2
4
6
8
10


Atrorosin Q
HMBC
DMSO-d6
F1 [ppm]
50
100
150
200
8
6
4
2
F2 [ppm]

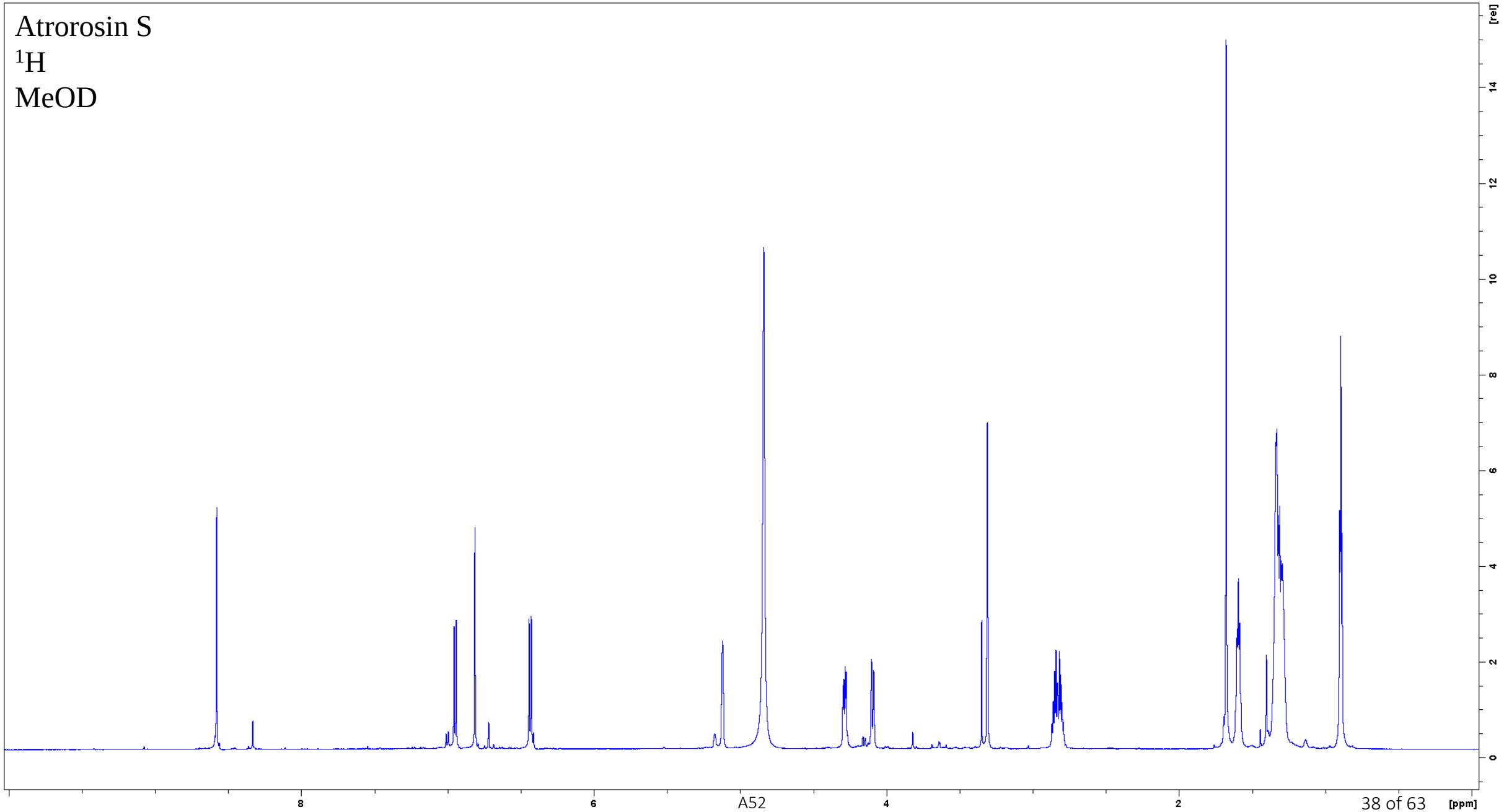

Atrorosin R
1H
MeOD
[rel]
3.0
2.5
2.0
1.5
1.0
0.5
-0.0
8
6
4
2
[ppm]
A50
36 of 63

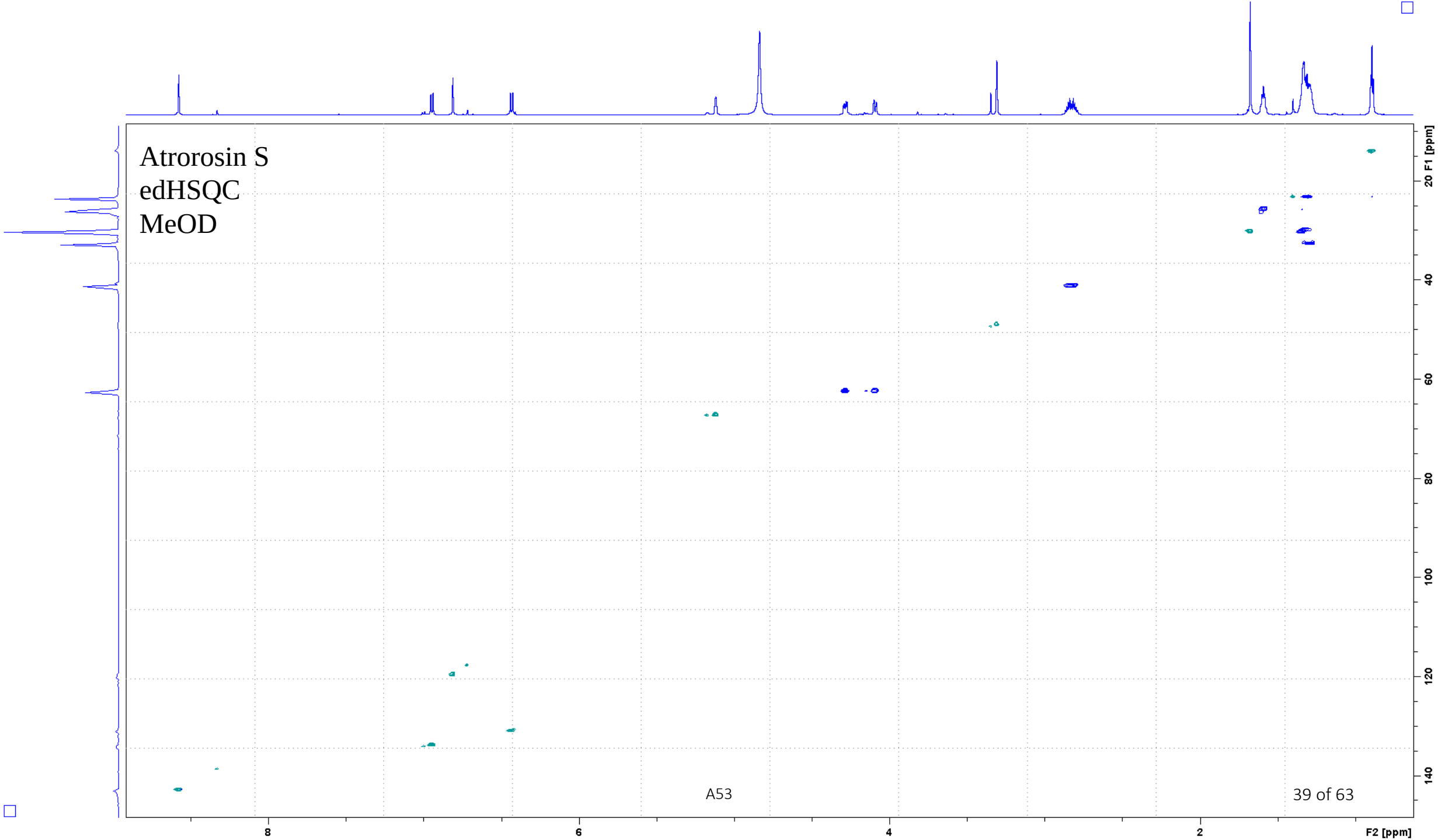

Atrorosin R
edHSQC
MeOD
A51
37 of 63
F1 [ppm]
F2 [ppm]
20
40
60
80
100
120
140
8
6
4
2

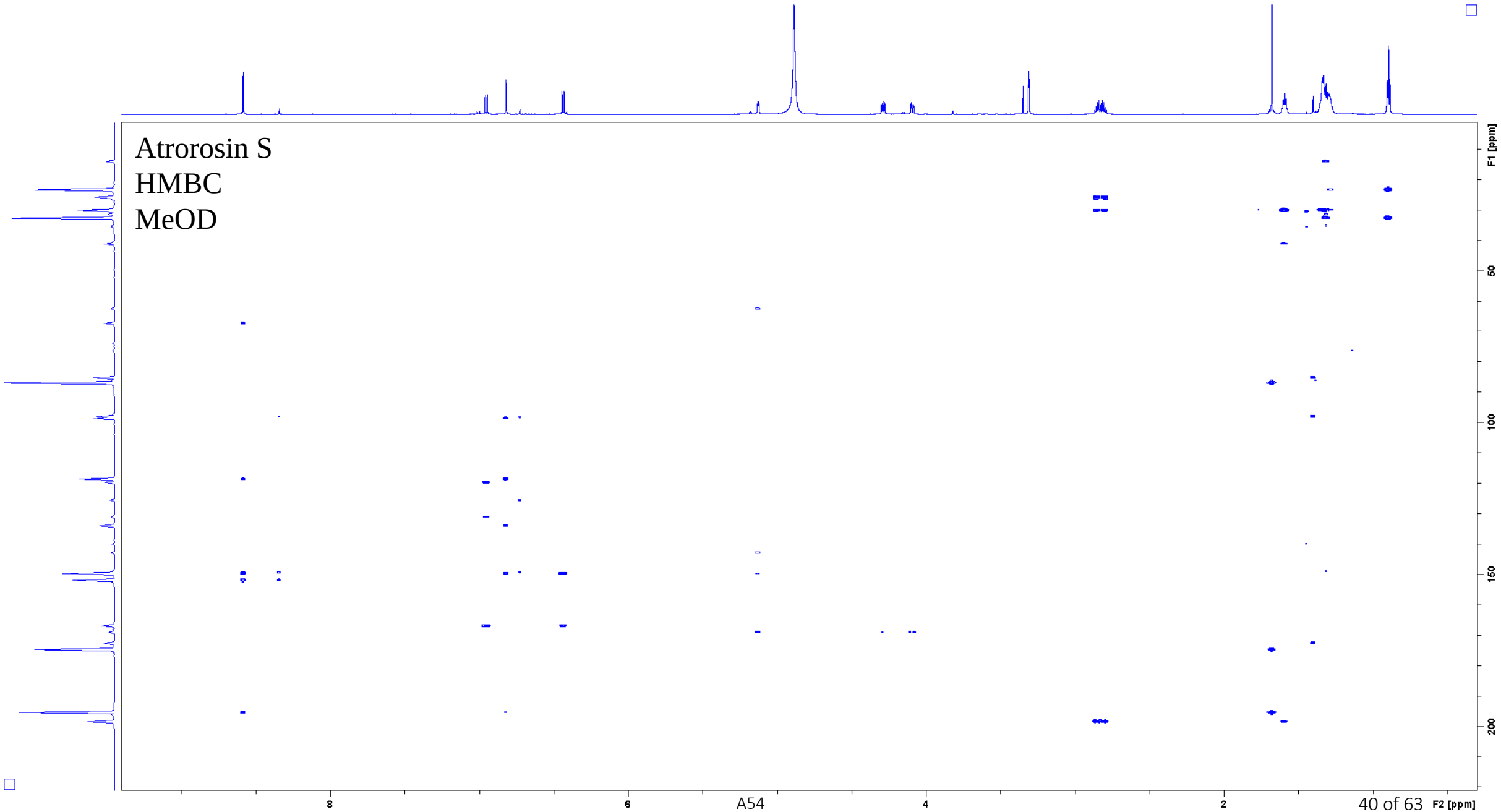
Atrorosin S

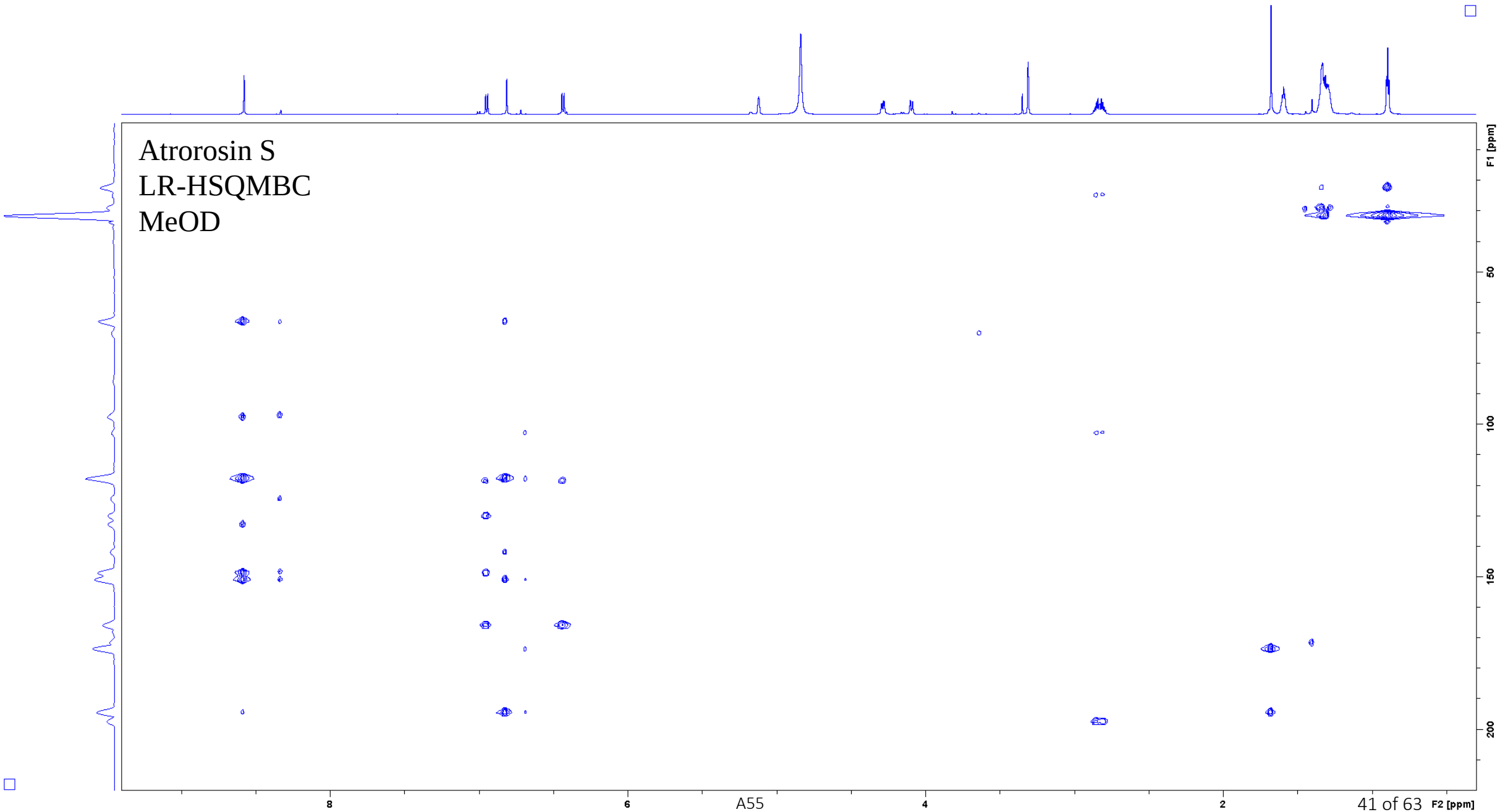
^{1}H

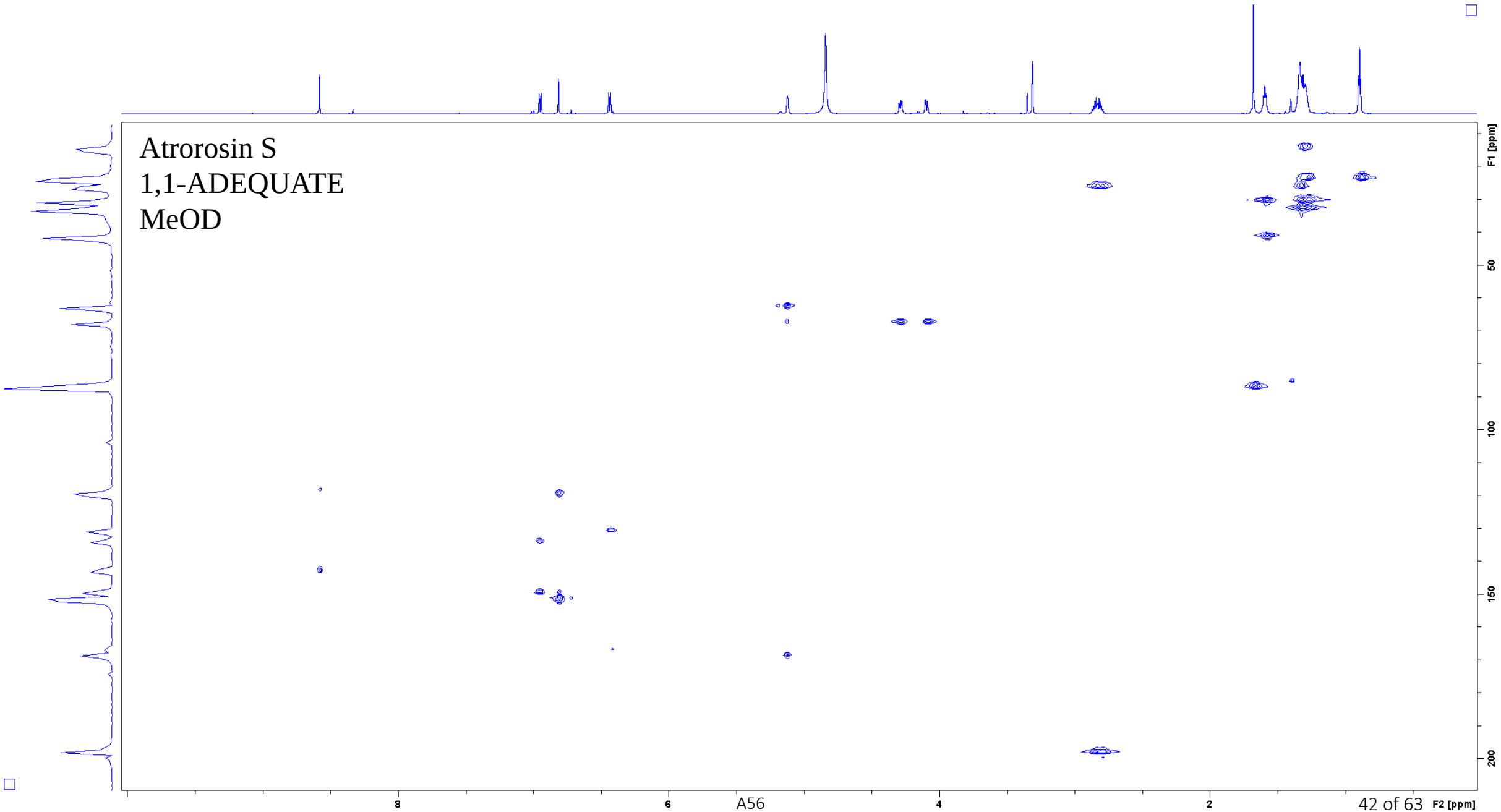
MeOD

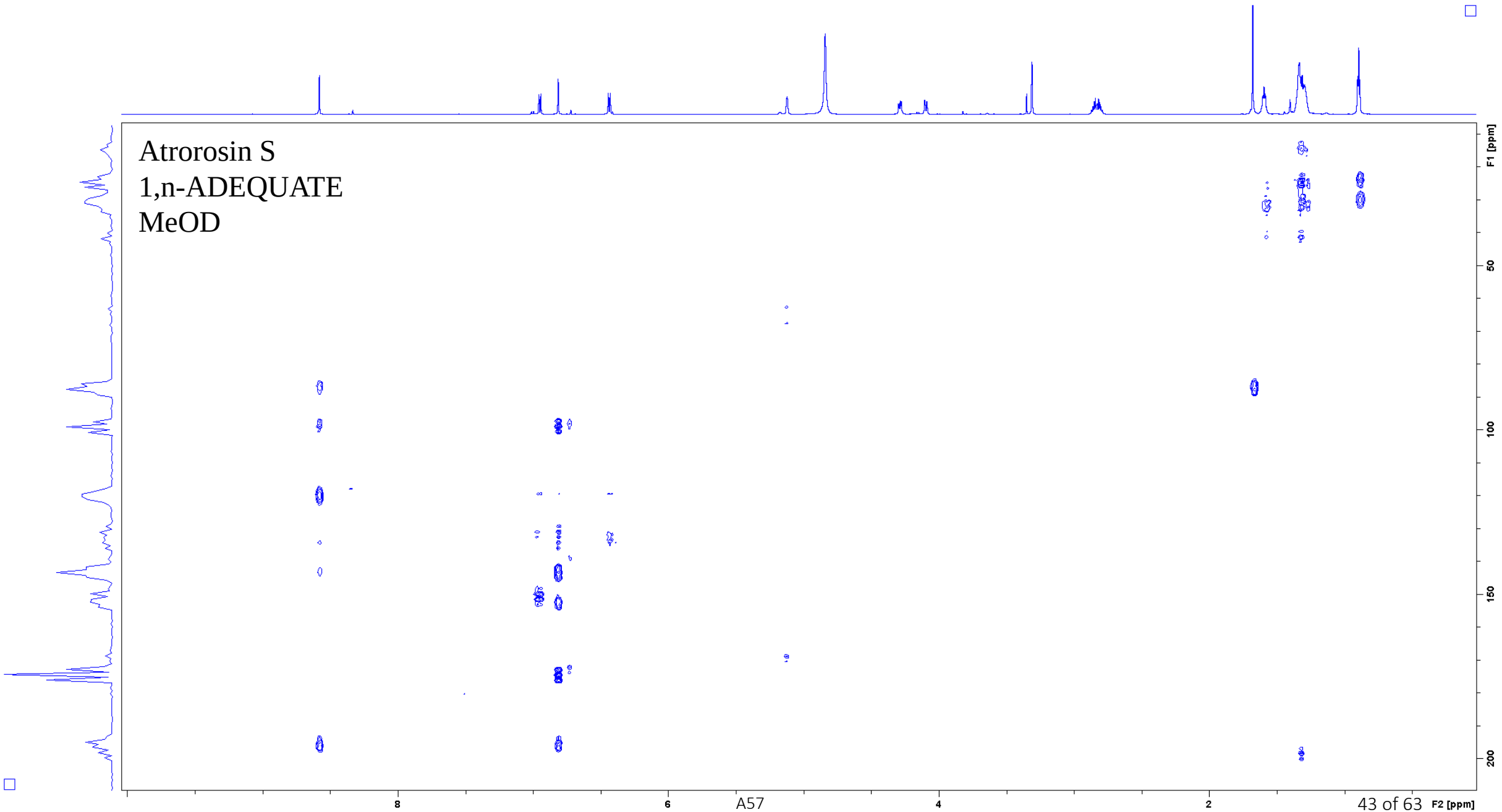


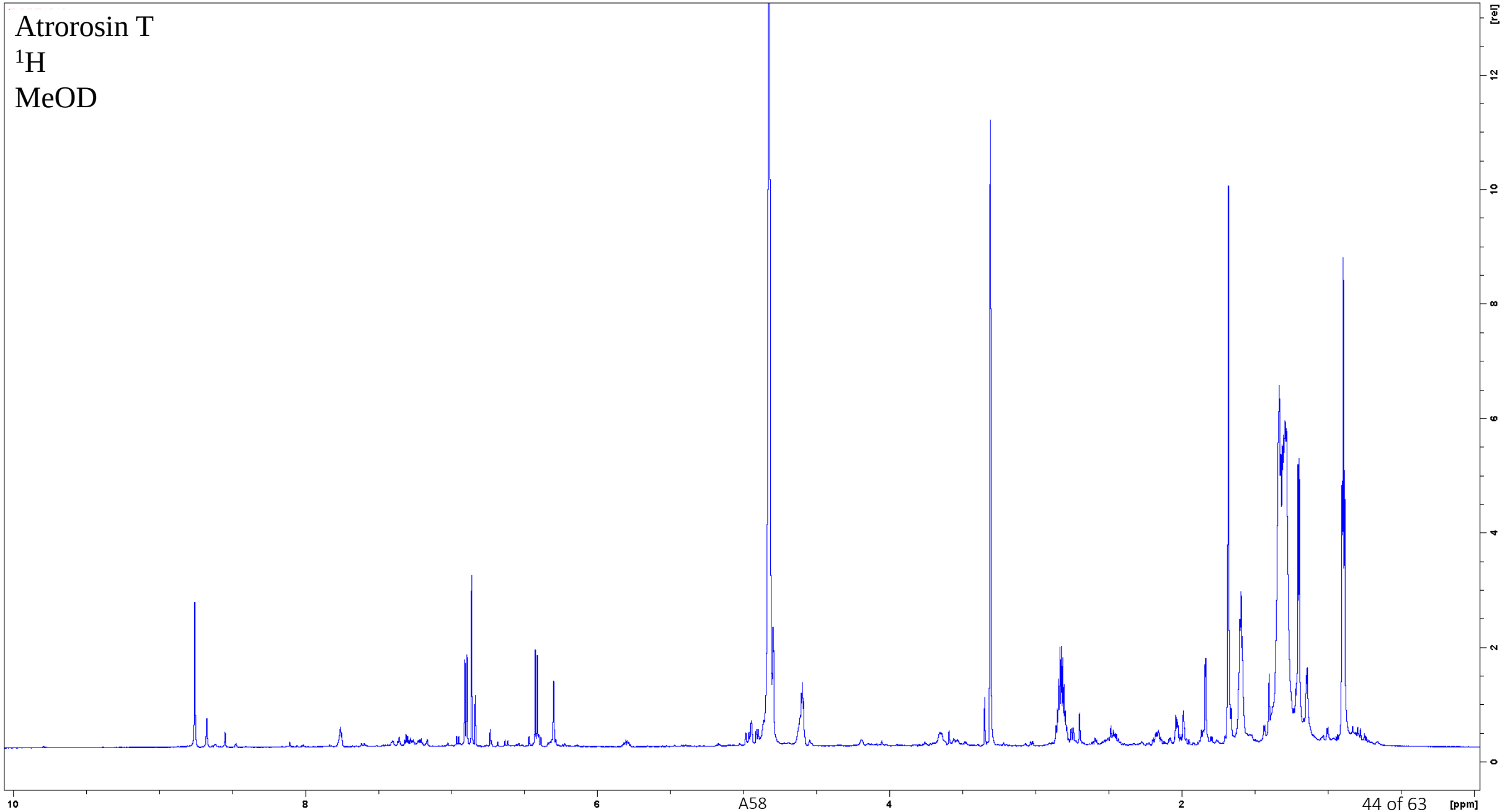

Atrorosin S
edHSQC
MeOD
A53
39 of 63
F2 [ppm]
F1 [ppm]
8
6
4
2
20
40
60
80
100
120
140

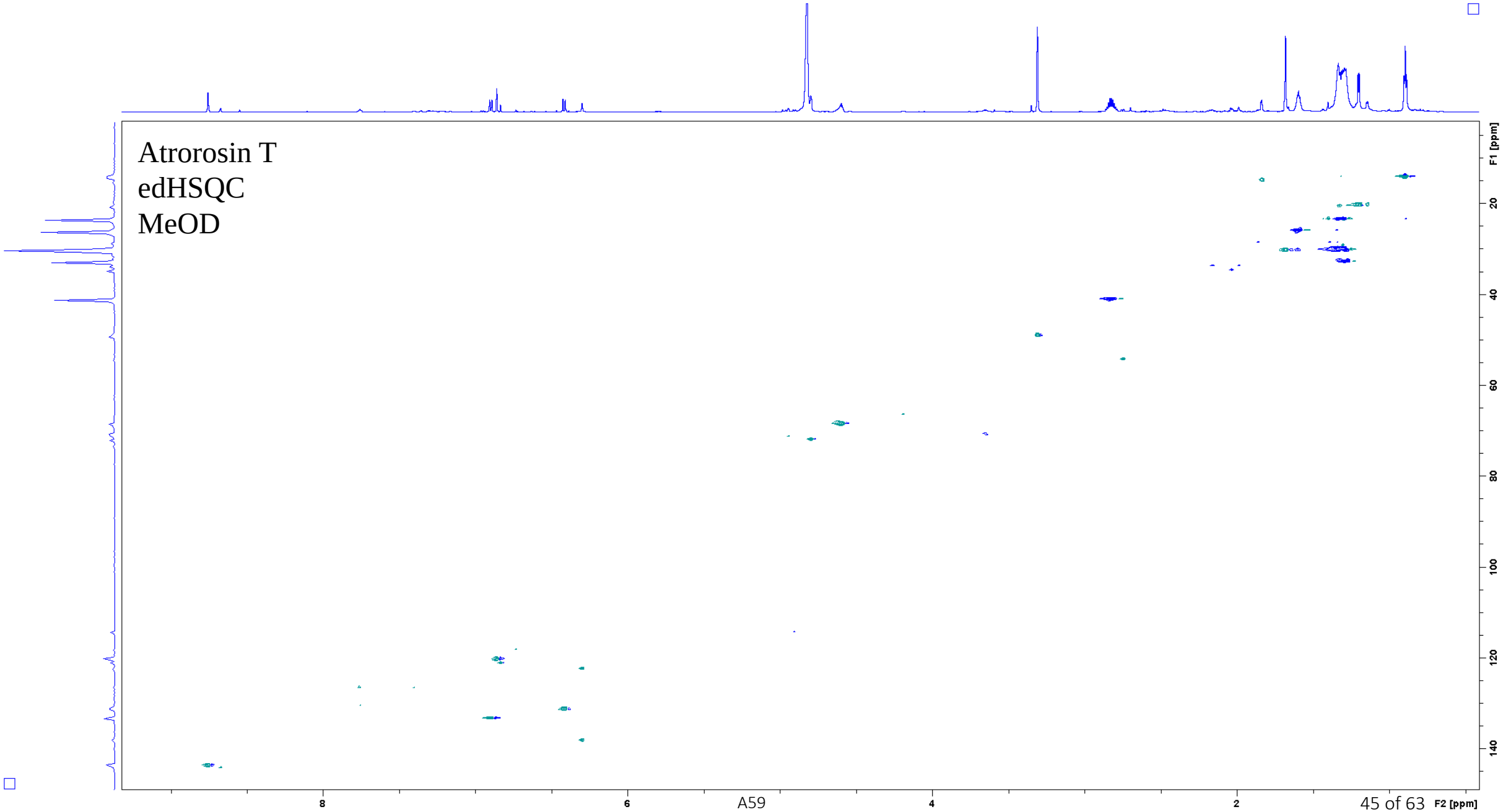

Atrorosin S
HMBC
MeOD
F1 [ppm]
50
100
150
200
8
6
4
2
F2 [ppm]

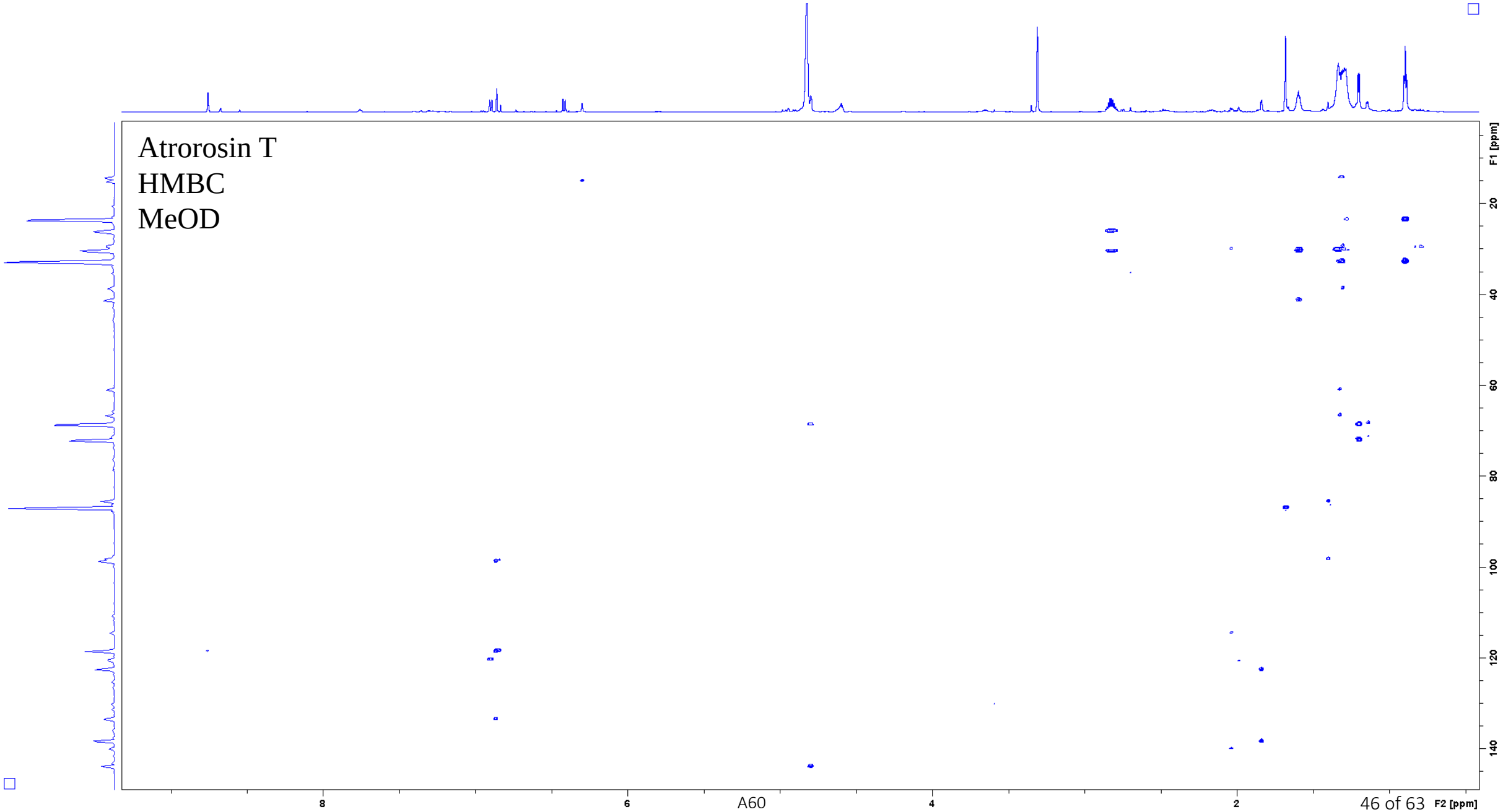

Atrorosin S
LR-HSQMBC
MeOD
F1 [ppm]
50
100
150
200
8
6
4
2
F2 [ppm]

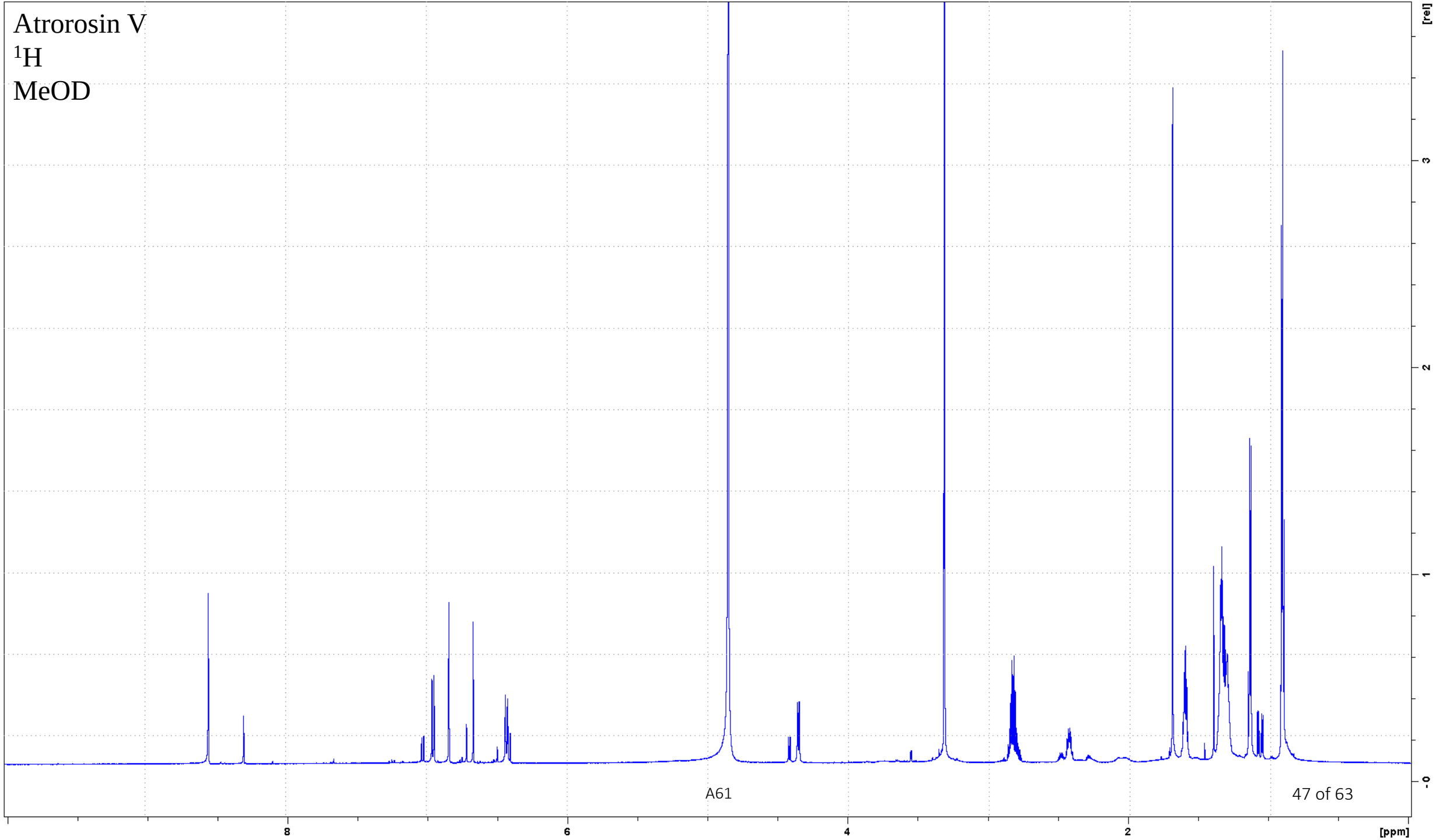

Atrorosin S
1,1-ADEQUATE
MeOD
F1 [ppm]
50
100
150
200
8
6
4
2
F2 [ppm]

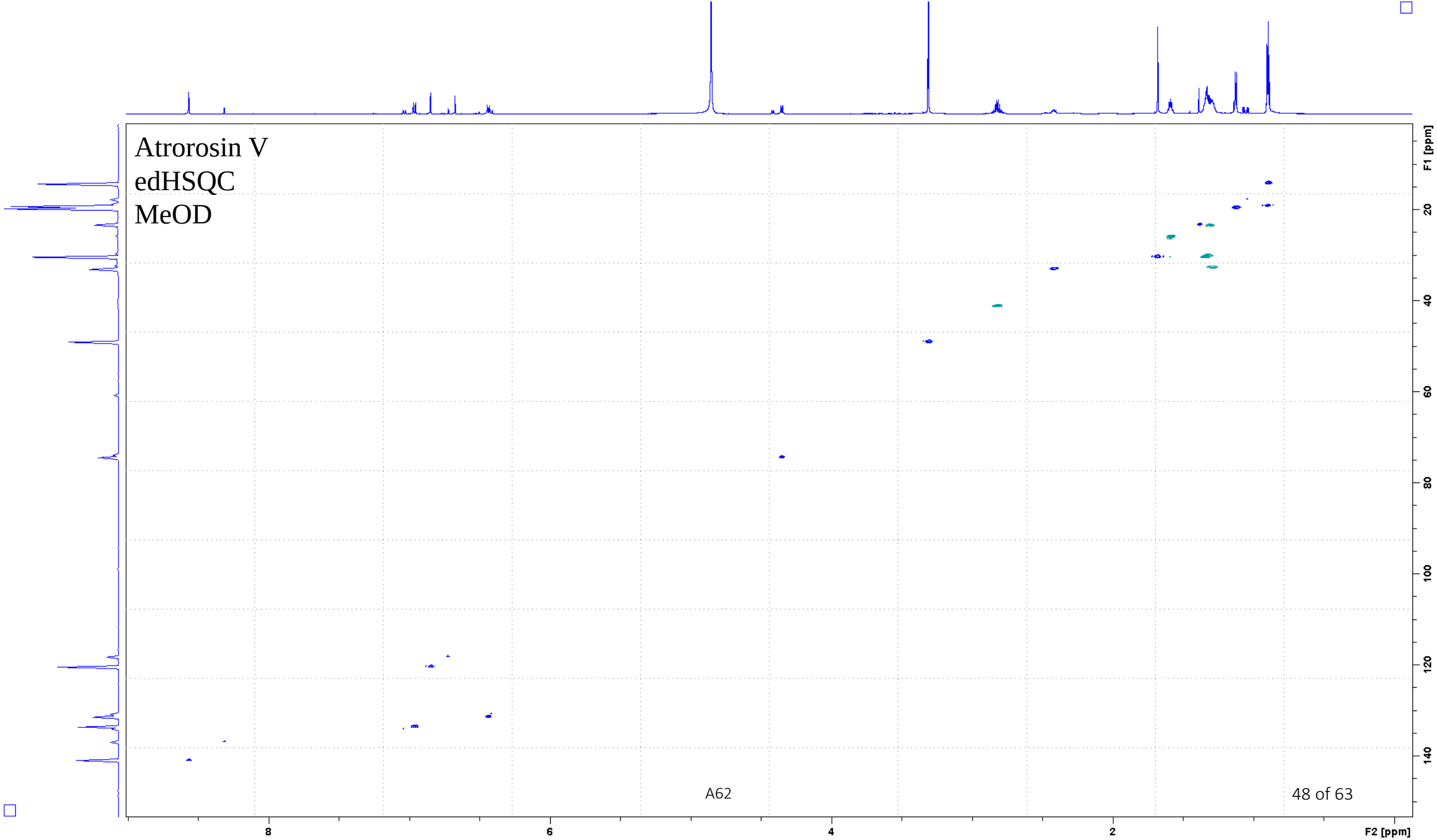

Atrorosin S
1,n-ADEQUATE
MeOD
F1 [ppm]
50
100
150
200
8
6
4
2
F2 [ppm]


Atrorosin T
1H
MeOD
[rel]
0
2
4
6
8
10
12
10
8
6
4
2
[ppm]


Atrorosin T
edHSQC
MeOD
F1 [ppm]
20
40
60
80
100
120
140
8
6
4
2
F2 [ppm]


Atrorosin T
HMBC
MeOD
F1 [ppm]
20
40
60
80
100
120
140
8
6
4
2
F2 [ppm]

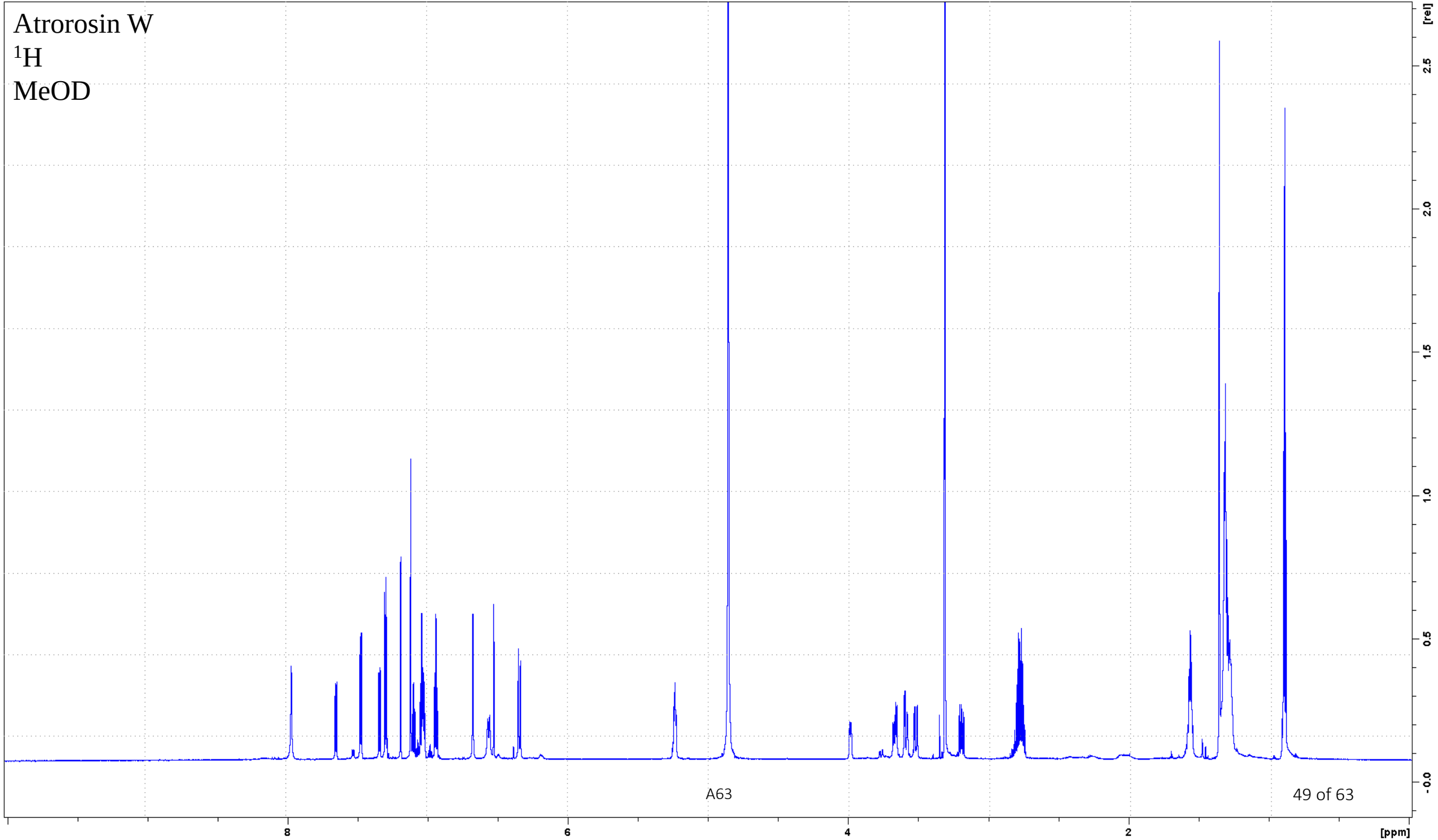

Atrorosin V
1H
MeOD
A61
47 of 63
8
6
4
2
[ppm]
[rel]
3
2
1
-0


Atrorosin V
edHSQC
MeOD
A62
48 of 63
F1 [ppm]
F2 [ppm]
20
40
60
80
100
120
140
8
6
4
2

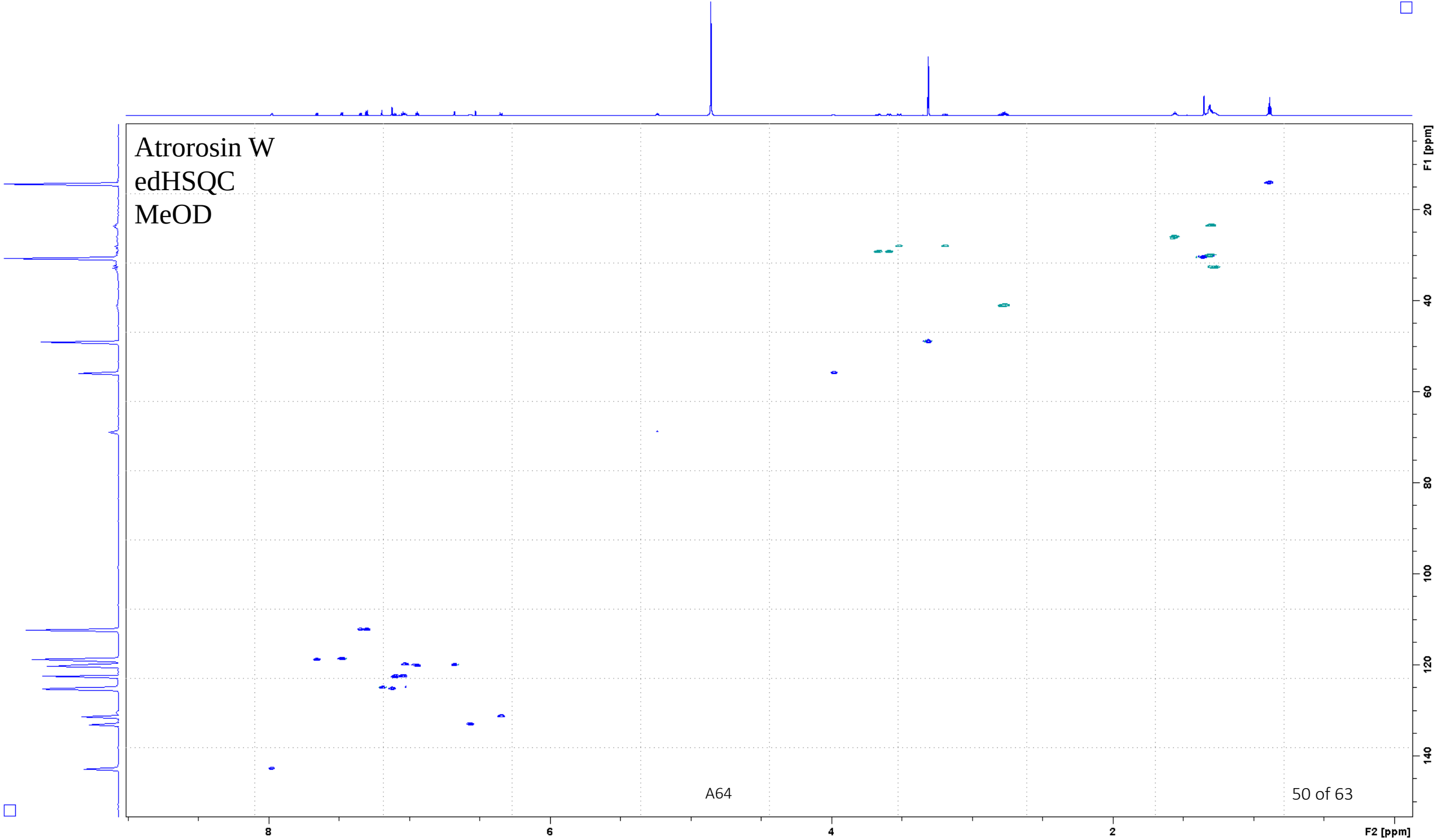
Atrorosin W

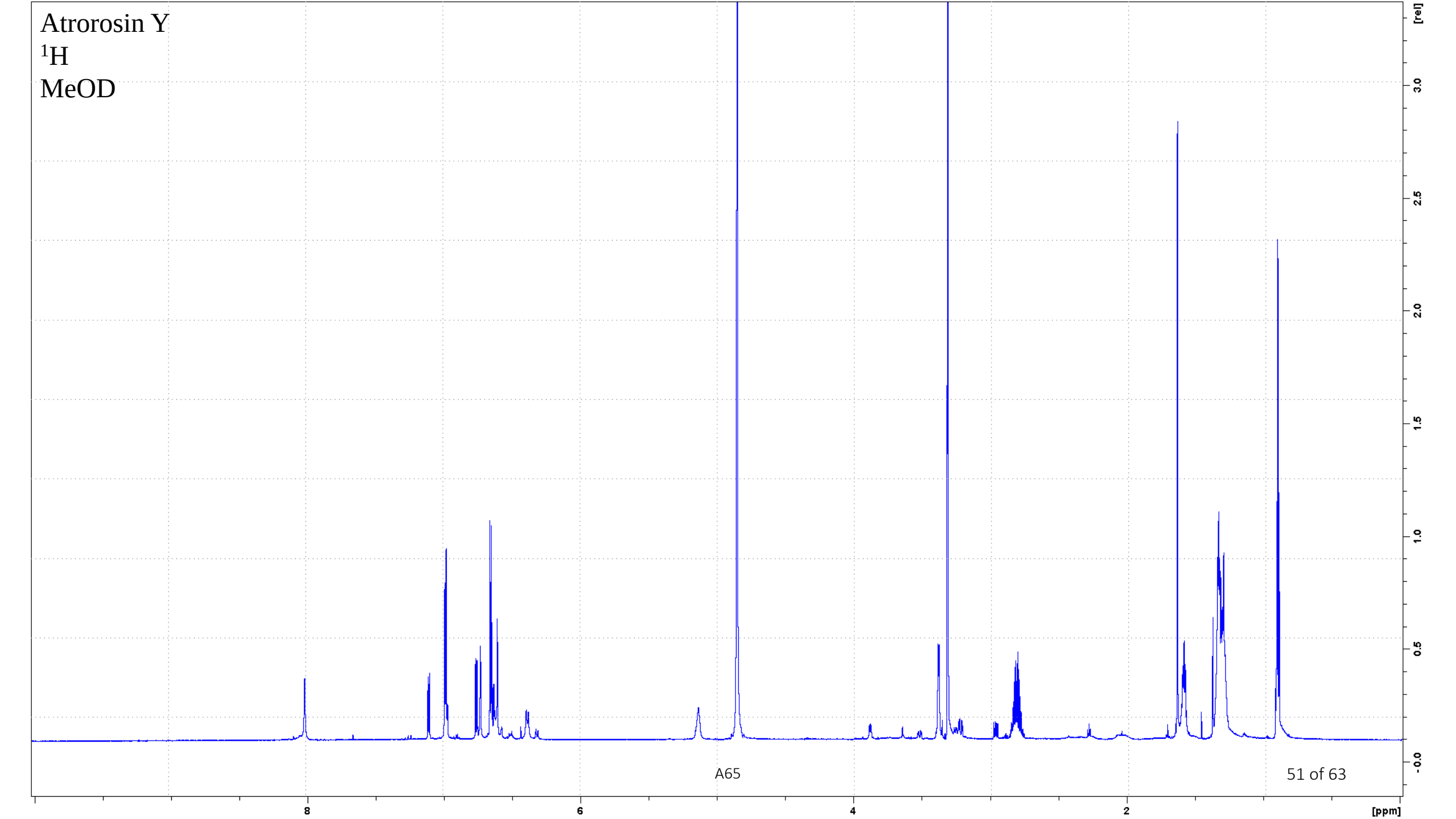
^{1}H

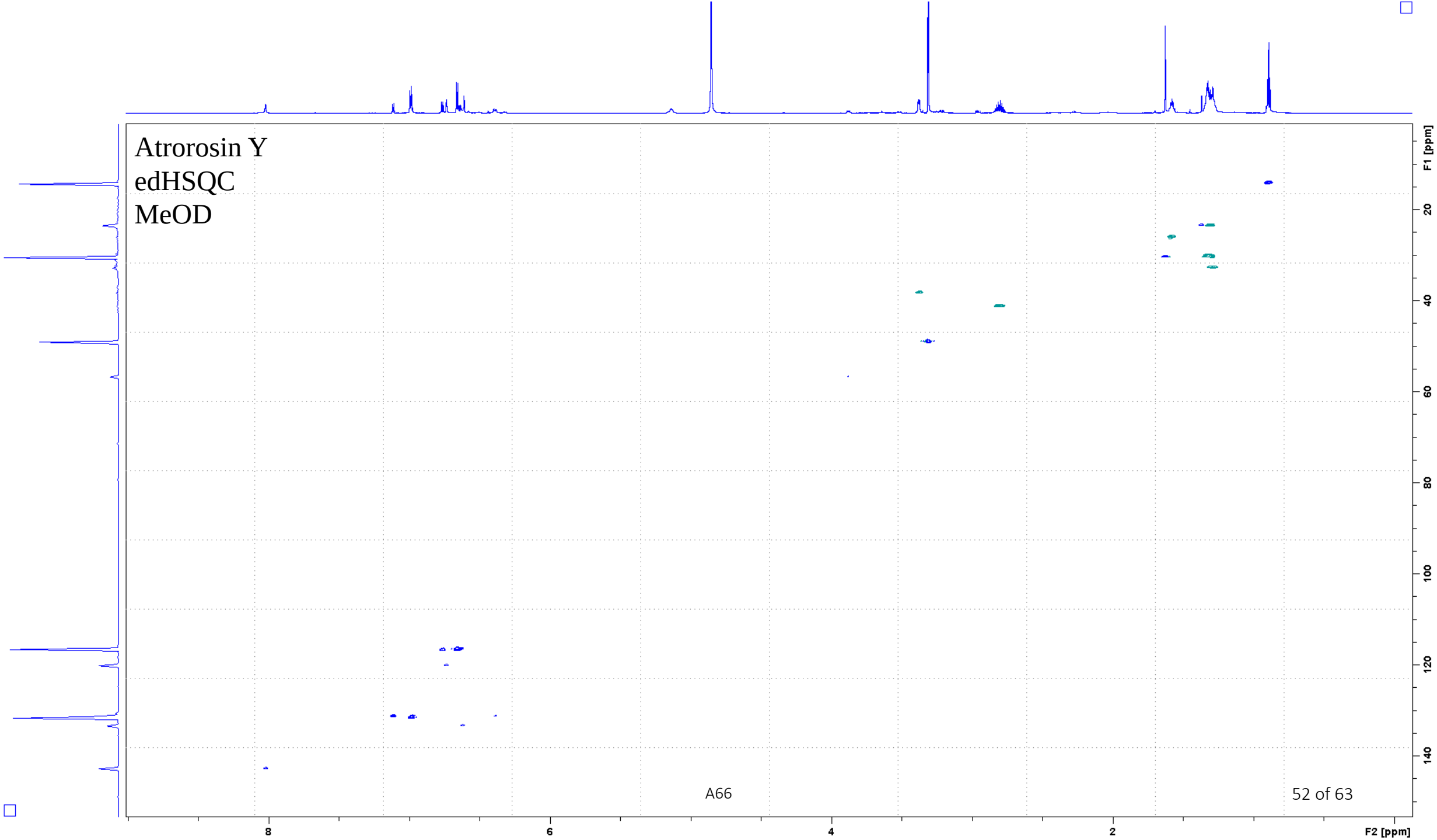
MeOD

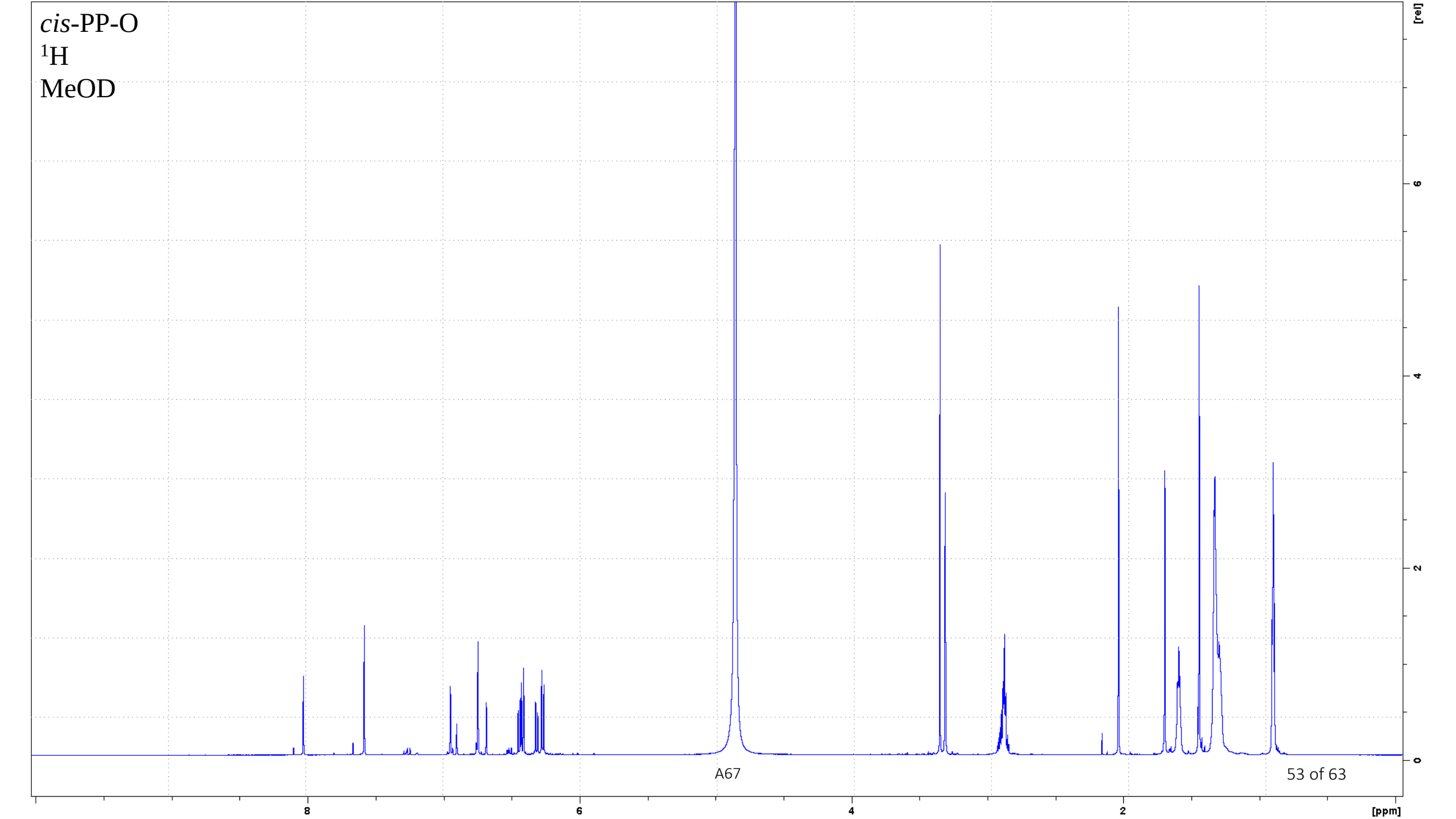


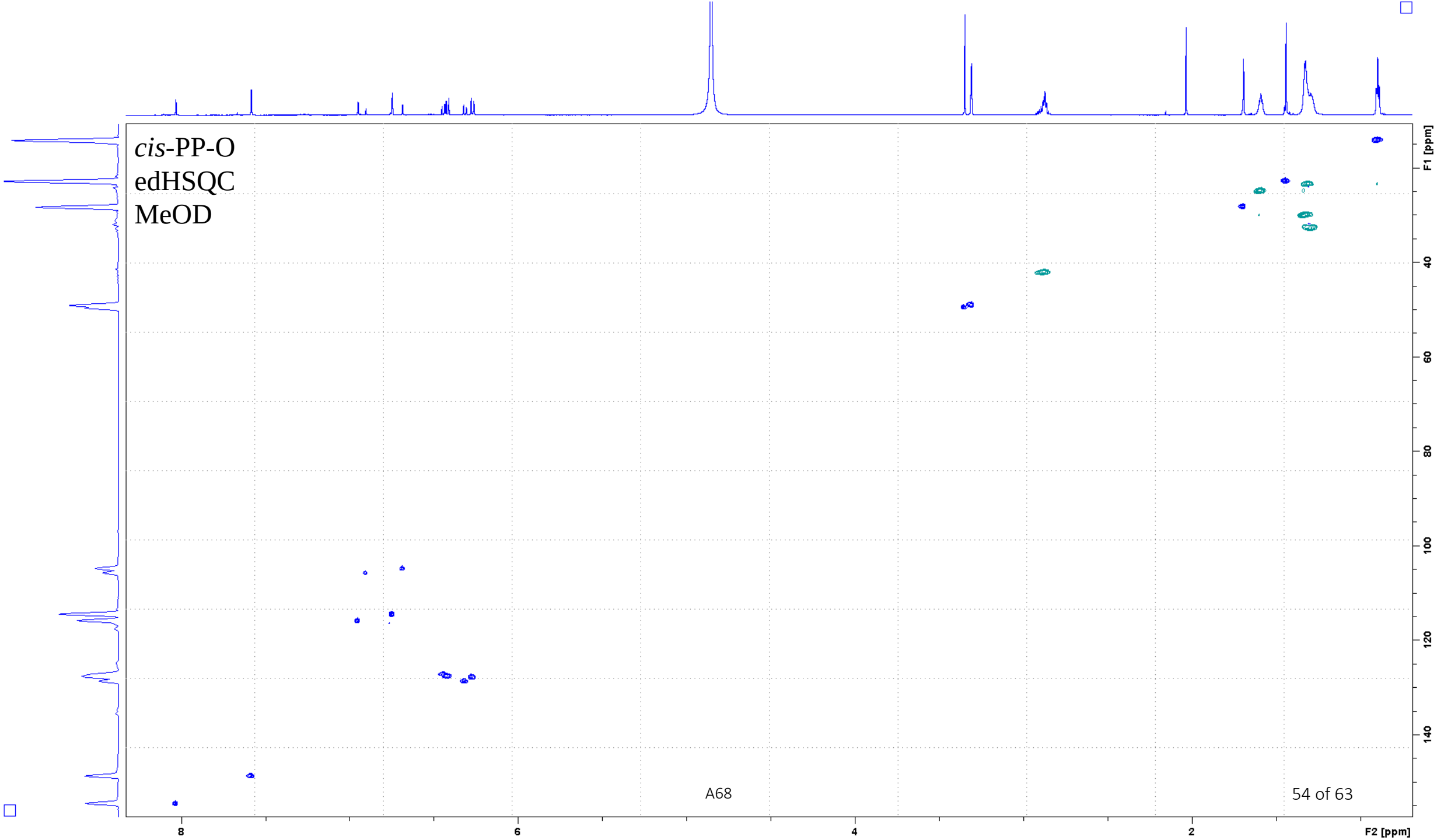
A63


Atrorosin W
edHSQC
MeOD
A64
50 of 63
F1 [ppm]
F2 [ppm]
20
40
60
80
100
120
140
8
6
4
2

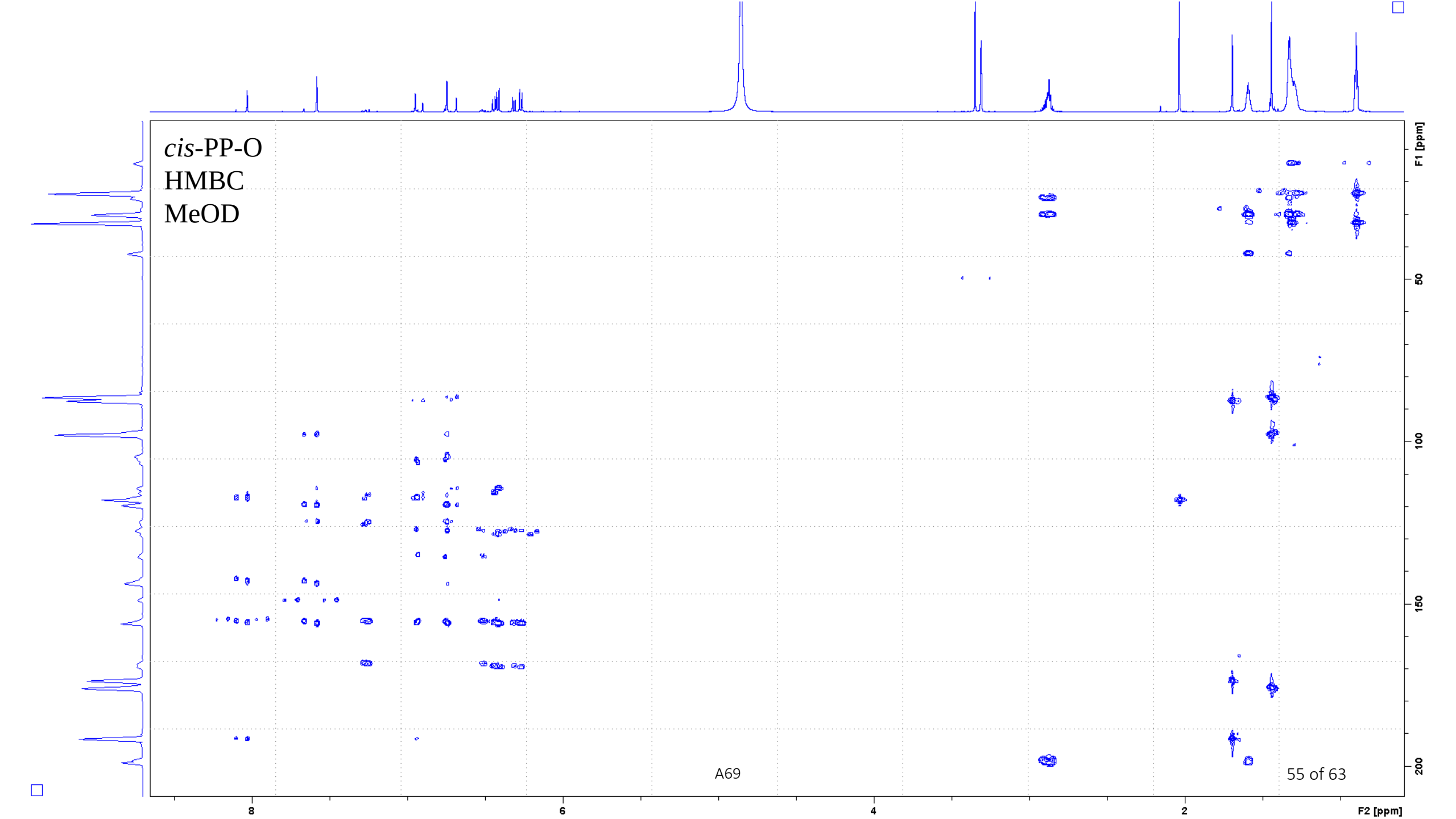

Atrorosin Y
1H
MeOD
A65
51 of 63
8
6
4
2
[ppm]
[rel]
3.0
2.5
2.0
1.5
1.0
0.5
-0.0

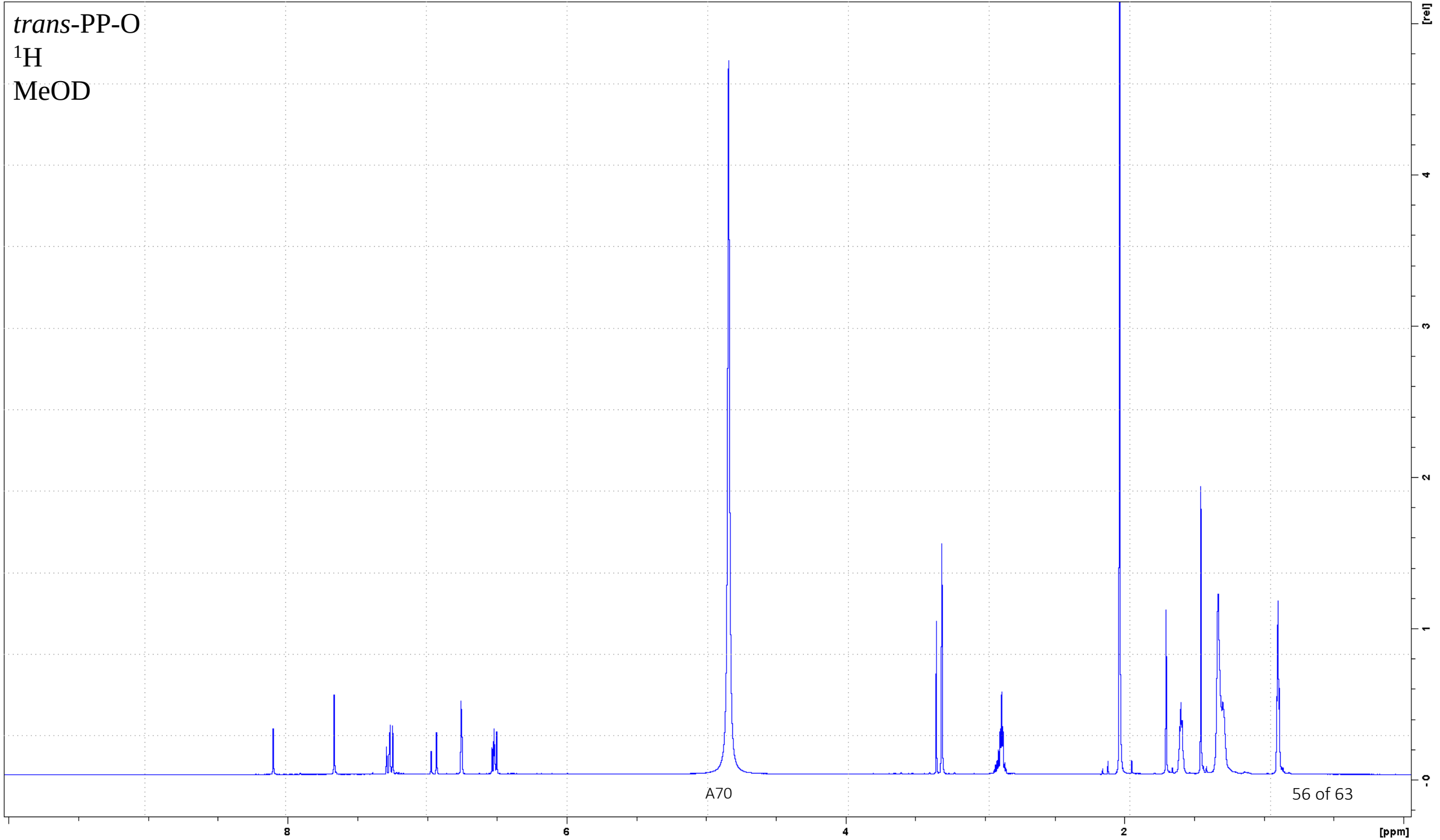

Atrorosin Y
edHSQC
MeOD
A66
52 of 63
F1 [ppm]
F2 [ppm]
20
40
60
80
100
120
140
8
6
4
2

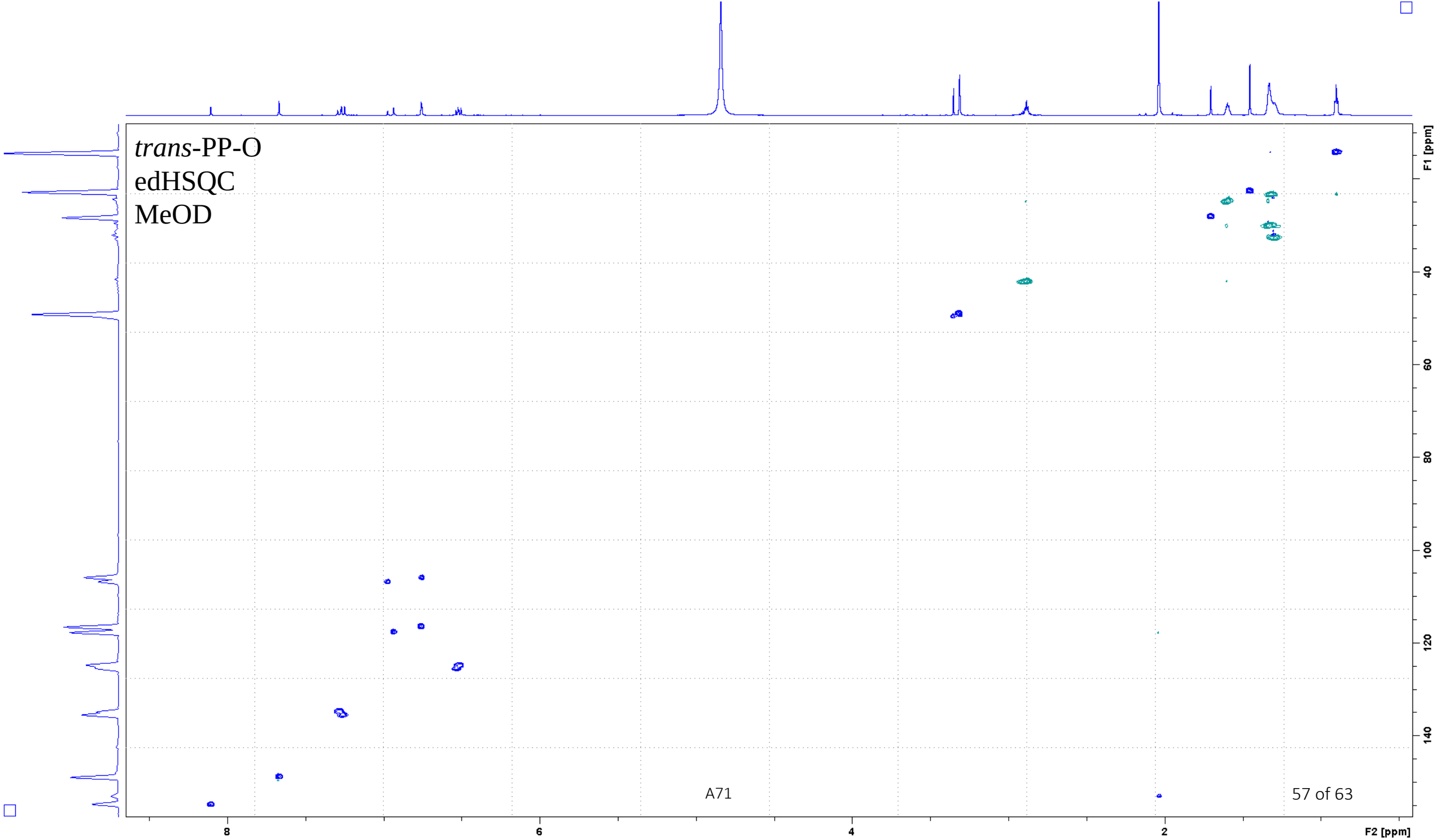

cis-PP-O
^{1}H
MeOD
[rel]
0
2
4
6
8
6
4
2
[ppm]
A67
53 of 63

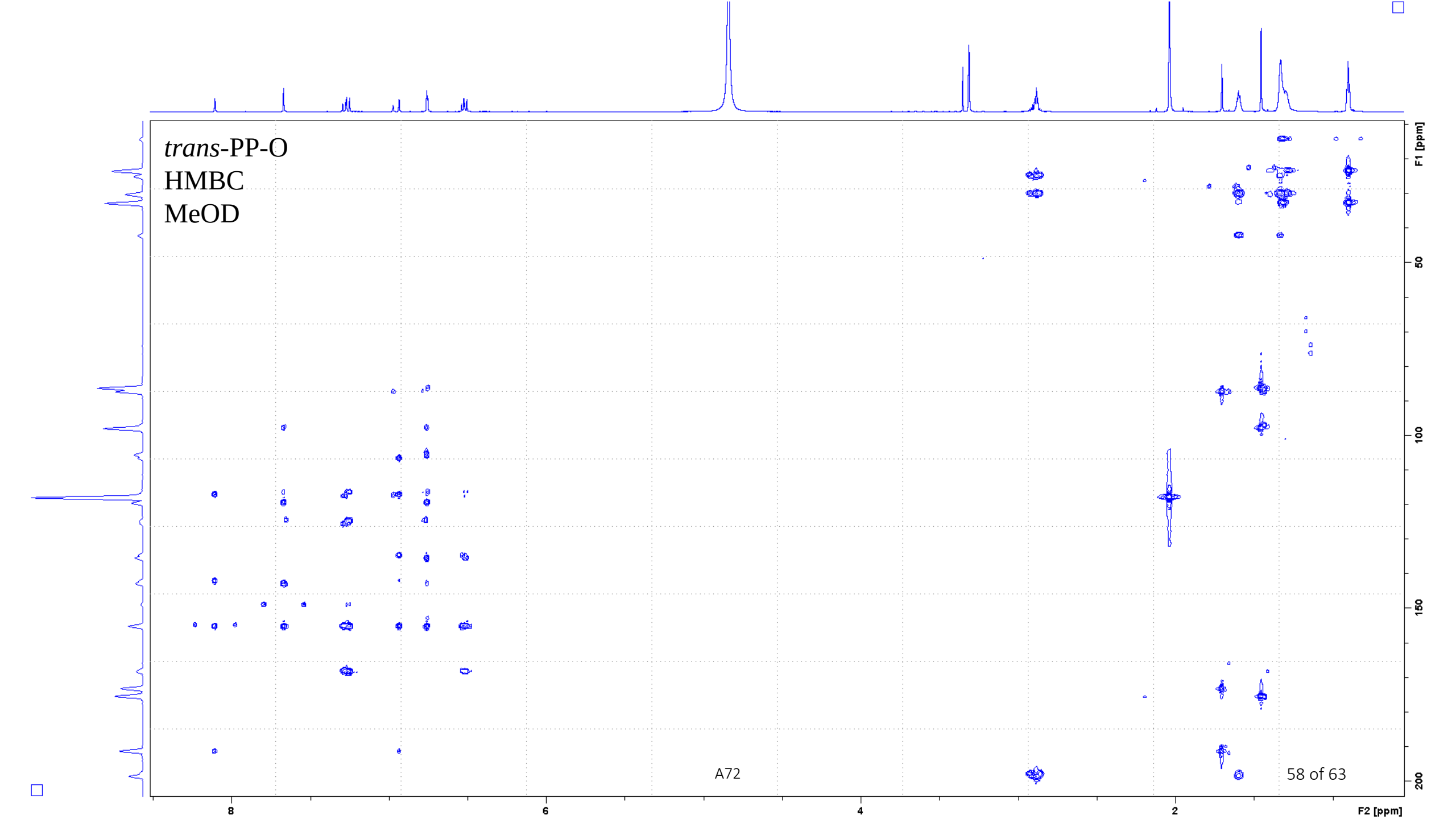


A68


cis-PP-O
HMBC
MeOD
F1 [ppm]
50
100
150
200
8
6
4
2
F2 [ppm]
A69
55 of 63


trans-PP-O
^{1}H
MeOD
[rel]
4
3
2
1
-0
8
6
4
2
[ppm]
A70
56 of 63

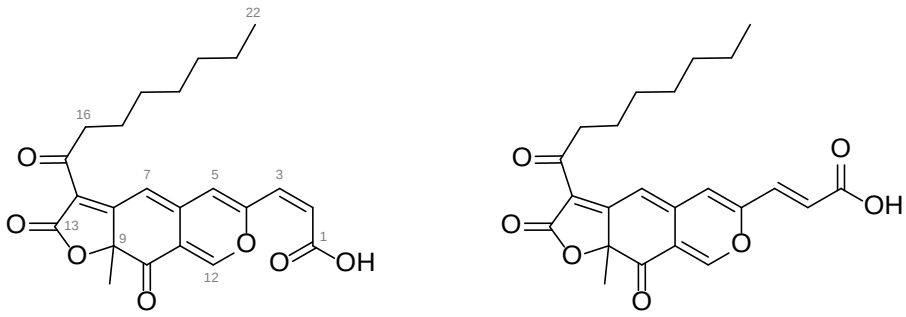

trans-PP-O
edHSQC
MeOD
F1 [ppm]
40
60
80
100
120
140
8
6
4
2
F2 [ppm]
A71
57 of 63


trans-PP-O
HMBC
MeOD
A72
58 of 63
F1 [ppm]
F2 [ppm]
50
100
150
200
8
6
4
2

S2. NMR data for PP-O isomers

Table S1. Proton and carbon shifts for cis-PP-O and trans-PP-O measured in CD_3OD.

	cis-PP-O		*trans*-PP-O	
#	δ_H	δ_C	δ_H	δ_C
1	-	169.3	-	168.3
2	6.27 (d, *J*=12.8 Hz)	128.6	6.5 (d *J*=15.7 Hz)	124.7
3	6.42 (d, J=12.8 Hz)	127.6	7.25 (d *J*=15.7 Hz)	135.5
4	-	155.7	-	155.3
5	6.95 (s)	115.8	6.93 (s)	117.6
6	-	142.8	-	142.3
7	6.69 (s)	104.8	6.75 (s)	105.8
8	-	Not observed	-	Not observed
9	-	86.4	-	86.4
9-CH3	1.69 (s)	28.2	1.71 (s)	28.1
10	-	191.8	-	191.4
11	-	117.2	-	117.1
12	8.03 (s)	154.6	8.11 (s)	154.8
13	-	173.5	-	173.4
14	-	119.7	-	119.5
15	-	198.9	-	198.7
16	2.88 (m)	42.1	2.88 (m)	42.1
17	1.59 (q, J=7 Hz)	24.8	1.59 (q, *J*=7 Hz)	24.8
18	1.31 (m)	23.4/30.1/32.7	1.31 (m)	23.4/30.1/32.7
19	1.31 (m)	23.4/30.1/32.7	1.31 (m)	23.4/30.1/32.7
20	1.31 (m)	23.4/30.1/32.7	1.31 (m)	23.4/30.1/32.7
21	1.31 (m)	23.4/30.1/32.7	1.31 (m)	23.4/30.1/32.7
22	0.9 (m)	14.2	0.9 (m)	14.1



cis-PP-O *trans*-PP-O

Figure S1. Structures of cis- and trans-PP-O.

S3. NMR tables for all atrorosins

Table S2, part 1

	Atrorosin A		Atrorosin C		Atrorosin D		Atrorosin E		Atrorosin F		Atrorosin G		Atrorosin H	
#	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C
1-OH	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1	-	-*	-	-*	-	167.1	-	-*	-	-*	-	-*	-	167.6
2	6.45 (d, 11.7)	131.4	6.72 (s)	99.0	6.43 (d, 11.7)	131.0	6.42 (d, 11.6)	131.2	6.37 (d, 11.8)	131.2	6.42 (d, 11.9)	131.1	6.46 (d, 11.7)	131.7
3	6.93 (d, 11.7)	133.6	7.00 (s)	119.7	7.07 (d, 11.7)	133.9	6.94 (d, 11.6)	134.1	6.59 (d, 11.6)	133.4	6.84 (d, 11.9)	133.2	6.89 (d, 11.7)	133.4
4	-	-*	-	-*	-	150.4	-	-*	-	-*	-	-*	-	149.2
5	6.81 (s)	119.9	6.92 (s)	117.9	6.82 (s)	120.1	6.80 (s)	119.9	6.71 (s)	120.1	6.88 (s)	120.2	6.78 (s)	120.6
6	-	-*	-	-*	-	152.1	-	-*	-	-*	-	-*	-	151.4
7	6.68 (s)	117.7	6.83 (s)	116.6	6.69 (s)	117.9	6.67 (s)	117.8	6.55 (s)	118.1	6.75 (s)	118.1	6.63 (s)	118.0
8	-	-*	-	-*	-	Not observed	-	-*	-	-*	-	-*	-	Not observed
9	-	-*	-	-*	-	87.2	-	-*	-	-*	-	-*	-	86.9
9-CH3	1.67 (s)	30.2	1.69 (s)	30.2	1.66 (s)	30.1	1.69 (s)	30.2	1.61 (s)	30.3	1.68 (s)	30.2	1.66 (s)	30.1
10	-	-*	-	-*	-	195.4	-	-*	-	-*	-	-*	-	195.3
11	-	-*	-	-*	-	98.8	-	-*	-	-*	-	-*	-	98.9
12	8.31 (s)	140.9	8.21 (s)	145.6	8.38 (s)	142.3	8.3 (s)	141.2	8.08 (s)	142.4	8.28 (s)	144.3	8.13 (s)	141.5
13	-	-*	-	-*	-	174.6	-	-*	-	-*	-	-*	-	174.6
14	-	-*	-	-*	-	126.1	-	-*	-	-*	-	-*	-	125.3
15	-	-*	-	-*	-	198.9	-	-*	-	-*	-	-*	-	198.6
16	2.82 (m)	41.1	2.82 (m)	41.1	2.82 (m)	41.0	2.83 (m)	41.3	2.81 (m)	41.1	2.82 (m)	41.2	2.82 (m)	41.2
17	1.59 (quin, 7)	25.9	1.59 (quin, 7.2)	26.1	1.59 (quin, 6.8)	25.9	1.59 (quin, 6.8)	25.8	1.57 (quin, 7.3)	25.9	1.59 (quin, 7.2)	25.9	1.58 (quin, 6.9)	25.8
18	1.31 (m)	23.3/30.1/32.6	1.31 (m)	23.4/26.9/30.3/32.7	1.31 (m)	23.3/30.1/32.6	1.31 (m)	23.4/30.2/32.7	1.31 (m)	23.4/30.2/32.7	1.31 (m)	23.3/30.1/32.7	1.31 (m)	23.4/30.2/32.7
19	1.31 (m)	23.3/30.1/32.6	1.31 (m)	23.4/26.9/30.3/32.7	1.31 (m)	23.3/30.1/32.6	1.31 (m)	23.4/30.2/32.7	1.31 (m)	23.4/30.2/32.7	1.31 (m)	23.3/30.1/32.7	1.31 (m)	23.4/30.2/32.7
20	1.31 (m)	23.3/30.1/32.6	1.31 (m)	23.4/26.9/30.3/32.7	1.31 (m)	23.3/30.1/32.6	1.31 (m)	23.4/30.2/32.7	1.31 (m)	23.4/30.2/32.7	1.31 (m)	23.3/30.1/32.7	1.31 (m)	23.4/30.2/32.7
21	1.31 (m)	23.3/30.1/32.6	1.31 (m)	23.4/26.9/30.3/32.7	1.31 (m)	23.3/30.1/32.6	1.31 (m)	23.4/30.2/32.7	1.31 (m)	23.4/30.2/32.7	1.31 (m)	23.3/30.1/32.7	1.31 (m)	23.4/30.2/32.7
22	0.89 (t, 7)	14.1	0.89 (t, 7)	14.2	0.89 (t, 7)	14.0	0.89 (t, 7)	14.1	0.89 (t, 7.1)	14.0	0.89 (t, 7)	140.2	0.89 (t, 7)	14.1
1'-OH	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1'	-	-*	-	-*	-	173.1	-	-*	-	-*	-	-*	-	170.7
2'	5.02 (quar, 7.2)	63.1	5.36 (m)	66.8	5.29 (broad t, n/a)	Not observed	5.15 (dd, 5.8, 9.5)	65.3	5.21 (dd, 8.5, 5.5)	68.2	4.77 (s)	57.4	5.15 (broad t, n/a)	66.3
3'	1.76 (d, 7.2)	18.2	3.62 (d, 13.4) / 3.49 (13.3, 5)	25.8	3.31 (m) / 3.14 (dd, 8, 17.6)	37.5	2.56/2.37 (m)	28.1	3.53/3.45	38.9	/	/	3.67 (dd, 6.3, 15.8) / 3.59 (dd, 15.8, 9)	28.1
4'	/	/	/	/	-	170.9	2.45 (td, 6.6, 2.3)	30.2	-	-	/	/	-	130.1
5'	/	/	/	/	/	/	-	-*	7.18 (m)	130.3	/	/	6.63 (s)	118.7
6'	/	/	/	/	/	/	/	/	7.25 (m)	129.7	/	/	8.72 (s)	135.4
7'	/	/	/	/	/	/	/	/	7.22 (m)	128.4	/	/	-	-
8'	/	/	/	/	/	/	/	/	/	/	/	/	/	/
9'	/	/	/	/	/	/	/	/	/	/	/	/	/	/
10'	/	/	/	/	/	/	/	/	/	/	/	/	/	/
11'	/	/	/	/	/	/	/	/	/	/	/	/	/	/
12'	/	/	/	/	/	/	/	/	/	/	/	/	/	/
5'-NH2	/	/	/	/	/	/	/	/	/	/	/	/	/	/

* Quaternary carbons not measured for semi-synthetic compounds

Table S2, part 2

	Atrorosin I		Atrorosin K1		Atrorosin K2		Atrorosin L		Atrorosin M		Atrorosin N		Atrorosin Q†	
#	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C
1-OH	-	-	-	-	-	-	-	-	-	-	-	-	12.5-14	-
1	-	-*	-	-*	-	-*	-	167.1	-	166.8	-	-*	-	165.6
2	6.43 (d, 11.8)	131.2	6.46 (d, 11.7)	131.3	6.46 (11.8)	130.8	6.46 (d, 11.6)	131.1	6.46 (d, 11.7)	131.2	6.43 (d, 11.7)	131.2	6.37 (d, 11.8)	130.0
3	6.97 (d, 11.8)	133.8	6.93 (d, 11.7)	133.8	7.1 (d, 11.9)	133.9	6.97 (d, 11.6)	134.0	7.01 (d, 11.6)	134.2	7.05 (d, 11.8)	133.6	6.94 (d, 11.8)	132.2
4	-	-*	-	-*	-	-*	-	149.5	-	149.8	-	-*	-	147.7
5	6.84 (s)	120.4	6.82 (s)	120.2	6.86 (s)	120.9	6.81 (s)	120.1	6.81 (s)	120.0	6.83 (s)	120.3	6.89 (s)	118.6
6	-	-*	-	-*	-	-*	-	151.6	-	151.8	-	-*	-	149.4
7	6.71 (s)	118.2	6.69 (s)	117.9	6.77 (s)	119.0	6.69 (s)	117.9	6.7 (s)	117.9	6.69 (s)	118.1	6.59 (s)	96.4
8	-	-*	-	-*	-	-*	-	Not observed	-	Not observed	-	-*	-	172.6
9	-	-*	-	-*	-	-*	-	87.2	-	87.0	-	-*	-	84.9
9-CH3	1.67 (s)	30.2	1.67 (s)	30.1	1.67 (s)	30.3	1.68 (s)	30.2	1.68 (s)	30.2	1.65 (s)	30.2	1.59 (s)	29.6
10	-	-*	-	-*	-	-*	-	195.5	-	195.1	-	-*	-	194.0
11	-	-*	-	-*	-	-*	-	98.8	-	98.1	-	-*	-	117.2
12	8.54 (s)	140.9	8.27 (s)	141.1	8.31 (s)	143.2	8.29 (s)	141.0	8.31 (s)	141.4	8.37 (s)	Not observed	8.22 (s)	140.5
13	-	-*	-	-*	-	-*	-	174.8	-	175.0	-	-*	-	169.3
14	-	-*	-	-*	-	-*	-	125.9	-	126.1	-	-*	-	103.2
15	-	-*	-	-*	-	-*	-	198.9	-	198.6	-	-*	-	194.6
16	2.82 (m)	41.2	2.82 (m)	41.2	2.82 (m)	41.1	2.83 (m)	41.1	2.83 (m)	41.2	2.82 (m)	41.2	2.72 (t, 7.3)	39.6
17	1.59 (quin, 7.3)	25.8	1.59 (quin, 7.1)	25.8	1.59 (quin, 7.2)	25.9	1.59 (quin, 6.7)	25.7	1.59 (quin, 7.2)	25.9	1.59 (quin, 7.1)	25.9	1.48 (quin, 6.7)	24.0
18	1.31 (m)	23.3/30.1/32.7	1.31 (m)	23.3/30.1/32.6	1.31 (m)	23.3/30.1/32.6	1.32 (m)	23.3/30.1/32.7	1.31 (m)	23.4/30.1/32.7	1.31 (m)	23.3/30.1/32.7	1.23 (m)	22.1/28.6/31.2
19	1.31 (m)	23.3/30.1/32.7	1.31 (m)	23.3/30.1/32.6	1.31 (m)	23.3/30.1/32.6	1.32 (m)	23.3/30.1/32.7	1.31 (m)	23.4/30.1/32.7	1.31 (m)	23.3/30.1/32.7	1.23 (m)	22.1/28.6/31.2
20	1.31 (m)	23.3/30.1/32.7	1.31 (m)	23.3/30.1/32.6	1.31 (m)	23.3/30.1/32.6	1.32 (m)	23.3/30.1/32.7	1.31 (m)	23.4/30.1/32.7	1.31 (m)	23.3/30.1/32.7	1.23 (m)	22.1/28.6/31.2
21	1.31 (m)	23.3/30.1/32.7	1.31 (m)	23.3/30.1/32.6	1.31 (m)	23.3/30.1/32.6	1.32 (m)	23.3/30.1/32.7	1.31 (m)	23.4/30.1/32.7	1.31 (m)	23.3/30.1/32.7	1.23 (m)	22.1/28.6/31.2
22	0.89 (t, 7.1)	14.2	0.89 (t, 7.1)	14.1	0.89 (t, 7)	14.2	0.89 (t, 6.2)	14.0	0.9 (t, 7.2)	14.1	0.89 (t, 7)	14.1	0.85 (t, 7)	13.9
1'-OH	-	-	-	-	-	-	-	-	-	-	-	-	7.31 (s)	-
1'	-	-*	-	-*	4.09 (m)	55.9	-	171.4	-	170.7	-	-*	-	172.9
2'	4.46 (d, 9.8)	72.4	4.94 (dd, 9.3, 5.9)	66.3	1.85 (quin, 7.5)	30.2	4.94 (m)	64.9	5.21 (dd, 8.9, 5.2)	64.6	3.94 (m)	51.9	4.93 (bs)	47.8
3'	2.24 (m)	38.8	2.33/2.18 (m)	32.2	1.53/1.47 (m)	22.5	2.11/2.05 (m)	41.7	2.65/2.52 (m)	30.7	3.21 (dd, 16.3, 4.6) / 3.05 (dd, 16.4, 9.3)	38.6	2.33 (m)	26.9
4'	1.43/1.11 (m)	25.9	1.48/1.44 (m)	23.8	1.95/1.88 (m)	30.8	1.65 (m)	25.7	2.56/2.43 (m)	31.8	-	-*	2.13 (m)	30.6
5'	0.87 (t, 7.4)	10.8	1.68 (m)	27.7	3.94 (t, 6.3)	53.4	0.94 (d, 6.4)	22.3	2.06 (s)	14.5	/	/	-	170.3
6'	1.09 (d, 6.8)	15.6	2.91 (t, 7.6)	40.1	-	-*	/	/	/	/	/	/	-	/
7'	/	/	-	/	-	-	/	/	/	/	/	/	/	/
8'	/	/	/	/	/	/	/	/	/	/	/	/	/	/
9'	/	/	/	/	/	/	/	/	/	/	/	/	/	/
10'	/	/	/	/	/	/	/	/	/	/	/	/	/	/
11'	/	/	/	/	/	/	/	/	/	/	/	/	/	/
12'	/	/	/	/	/	/	/	/	/	/	/	/	/	/
5'-NH2	/	/	/	/	/	/	/	/	/	/	/	/	6.84 (s)	-

† Measured in DMSO-*d6*

* Quaternary carbons not measured for semi-synthetic compounds

Table S2, part 3

	Atrorosin R		Atrorosin S		Atrorosin T		Atrorosin V		Atrorosin W		Atrorosin Y	
#	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C
1-OH	-	-	-	-	-	-	-	-	-	-	-	-
1	-	-*	-	166.8	-	167.2	-	-*	-	-*	-	-*
2	6.74 (d, 11.6)	130.3	6.43 (d, 11.8)	130.9	6.41 (d, 11.7)	131.2	6.43 (d, 11.7)	131.4	6.34 (d, 11.8)	131.3	6.38 (d, 11.6)	131.3
3	6.45 (d, 11.7)	134.2	6.94 (d, 11.7)	133.8	6.89 (d, 11.7)	133.2	6.96 (d, 11.7)	133.5	6.56 (d, 11.7)	132.9	6.61 (m)	133.2
4	-	-*	-	149.7	-	150.2	-	-*	-	-*	-	-*
5	6.87 (s)	120.7	6.81 (s)	119.6	6.86 (s)	120.2	6.84 (s)	120.3	6.67 (s)	119.9	6.73 (s)	120.1
6	-	-*	-	151.8	-	152.2	-	-*	-	-*	-	-*
7	6.68 (s)	98.7	6.72 (s)	117.7	6.83 (s)	118.1	6.72 (s)	118.2	6.53 (s)	Not observed	6.58 (s)	118.2
8	-	-*	-	168.0	-	Not observed	-	-*	-	-*	-	-*
9	-	-*	-	86.8	-	86.9	-	-*	-	-*	-	-*
9-CH3	1.67 (s)	30.3	1.67 (s)	30.1	1.68 (s)	30.2	1.68 (s)	30.3	1.35 (s)	30.5	1.62 (s)	30.3
10	-	-*	-	195.4	-	196.0	-	-*	-	-*	-	-*
11	-	-*	-	98.8	-	98.6	-	-*	-	-*	-	-*
12	8.34 (s)	141.6	8.57 (s)	142.8	8.75 (s)	143.5	8.56 (s)	140.9	7.97 (s)	142.8	8.02 (s)	142.6
13	-	-*	-	174.7	-	174.4	-	-*	-	-*	-	-*
14	-	-*	-	125.4	-	125.2	-	-*	-	-*	-	-*
15	-	-*	-	198.5	-	198.5	-	-*	-	-*	-	-*
16	2.81 (m)	41.1	2.82 (m)	40.9	2.83 (m)	41.0	2.82 (m)	41.1	2.77 (m)	41.0	2.80 (m)	41.0
17	1.59 (quin, 7)	26	1.59 (quin, 7)	25.8	1.59 (quin, 6.5)	25.9	1.59 (quin, 7.1)	25.9	1.56 (quin, 7.2)	26.0	1.58 (quin, 7.2)	25.8
18	1.31 (m)	23.3/30.1/32.7	1.31 (m)	23.3/30.0/32.5	1.31 (m)	23.3/30.1/32.6	1.31 (m)	23.3/30.0/32.5	1.31 (m)	23.4/30.1/32.7	1.31 (m)	23.4/30.2/32.8
19	1.31 (m)	23.3/30.1/32.7	1.31 (m)	23.3/30.0/32.5	1.31 (m)	23.3/30.1/32.6	1.31 (m)	23.3/30.0/32.5	1.31 (m)	23.4/30.1/32.7	1.31 (m)	23.4/30.2/32.8
20	1.31 (m)	23.3/30.1/32.7	1.31 (m)	23.3/30.0/32.5	1.31 (m)	23.3/30.1/32.6	1.31 (m)	23.3/30.0/32.5	1.31 (m)	23.4/30.1/32.7	1.31 (m)	23.4/30.2/32.8
21	1.31 (m)	23.3/30.1/32.7	1.31 (m)	23.3/30.0/32.5	1.31 (m)	23.3/30.1/32.6	1.31 (m)	23.3/30.0/32.5	1.31 (m)	23.4/30.1/32.7	1.31 (m)	23.4/30.2/32.8
22	0.89 (t, 7)	14.1	0.89 (t, 6.2)	13.9	0.89 (t, 6.7)	14.0	0.89 (t, 7.3)	14.2	0.88 (t, 7.2)	14.1	0.89 (t, 7.1)	14.1
1'-OH	-	-	-	-	-	-	-	-	-	-	-	-
1'	-	-*	-	169.1	-	170.4	-	-*	-	-*	-	-*
2'	4.8 (dd, 8.8, 6.2)	68.4	5.12 (s)	67.3	4.59 (m)	68.3	4.35 (d, 10.2)	74.4	5.23 (m)	68.7	3.87 (dd, 8.4, 4.6)	56.6
3'	2.29/2.11 (m)	30.4	4.28 (dd, 12.2, 5.6) / 4.09 (d, 12.2)	62.5	4.79 (m)	71.8	2.42 (m)	33.0	3.52 (dd, 15.3, 4.3)) / 3.19 (dd, 15.3, 9.2)	27.9	3.22 (dd, 14.8, 4.6) / 2.96 (dd, 14.8, 8.5)	36.8
4'	1.68/1.61 (m)	26.5	/	/	1.2 (d, 6.3)	20.3	0.9 (m)	19.1	-	-*	-	-*
5'	3.18 (td, 6.8, 2)	41.4	/	/	/	/	1.12 (m)	19.6	7.12 (s)	125.1	6.99 (d, 8.5)	131.5
6'	/	/	/	/	/	/	/	/	-	-	6.65 (d, 8.5)	116.4
7'	/	/	/	/	/	/	/	/	-	-*	-	-*
8'	/	/	/	/	/	/	/	/	7.30 (d, 8.2)	112.2	/	/
9'	/	/	/	/	/	/	/	/	7.04 (m) / 6.94 (d, 7.4)	122.4 / 120.1	/	/
10'	/	/	/	/	/	/	/	/	6.94 (d, 7.4) / 7.04 (m)	120.1 / 122.4	/	/
11'	/	/	/	/	/	/	/	/	7.47 (d, 7.9)	118.7	/	/
12'	/	/	/	/	/	/	/	/	-	-*	/	/
5'-NH2	/	/	/	/	/	/	/	/	/	/	/	/

* Quaternary carbons not measured for semi-synthetic compounds

S4. Chemical derivatisation of PP-O

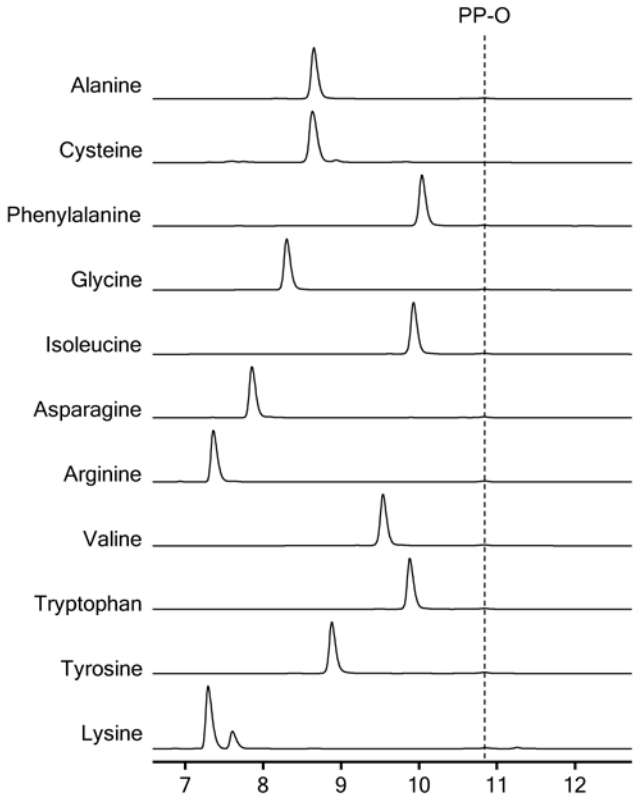


Figure S2. UV-VIS chromatograms at 520nm showing the amount of PP-O left after reaction with amino acids used for semi-synthesis of atrorosins

S5. Optical rotation values

Table S3. Measured optical rotation values for natural atrorosin E and S, as well as semi-synthetic versions with L- and D-amino acids incorporated.

	Atrorosin E	Atrorosin S
Natural	40 ± 11	56 ± 9
L-amino acid	29 ± 28	60 ± 7
D-amino acid	-32 ± 11	-49 ± 2

Appendix 2

Unique processes yielding pure azaphilones in *Talaromyces atroroseus*

Gerit Tolborg (genym@bio.dtu.dk)*, Anders S. R. Ødum (aodum@bio.dtu.dk) *, Thomas Isbrandt (tipse@bio.dtu.dk)*, Thomas Ostenfeld Larsen (tol@bio.dtu.dk) §* and Mhairi Workman (mwo@bio.dtu.dk)*

*DTU Bioengineering, Søltofts Plads, 2800 Lyngby

§ Address correspondence to this author

Abstract

Azaphilones are a class of fungal pigments, reported mostly in association with *Monascus* species. In Asian countries they are used as food colorants under the name of "red rice" and their production process is well described. One major limitation of current production techniques of azaphilones is that they always occur in a mixture of yellow, orange and red pigments. These mixtures are difficult to control and to quantify. This study has established a controlled and reproducible cultivation protocol to selectively tailor production of individual pigments during a submerged fermentation using another fungal species capable of producing azaphilone pigments, *Talaromyces atroroseus*, using single amino acids as the sole nitrogen source. The produced azaphilone pigments are called atrorosins and are amino acid derivatives of the known azaphilone pigment Penicilium purpurogenum- orange (PP-O), with the amino acid used as nitrogen source incorporated into the core skeleton of the azaphilone. This strategy was successfully demonstrated using 18 natural amino acids and the non-proteinogenic amino acid ornithine. Two cultivation methods for production of the pure serine derivative (atrorosin-S) have been further developed, with yields of 0.9 g/L being obtained. *In vitro* synthesis of atrorosins was demonstrated by incorporation of an amino acid precursor into PP-O. Chemical *in vitro* synthesis suggested the importance of the pH for amino acid incorporation. Yielding pure atrorosins through switching from potassium nitrate to single amino acids as nitrogen source allows for considerably easier downstream processing and thus further enhances the commercial relevance of azaphilone producing fungal cell factories.

Keywords: filamentous fungi, azaphilone pigments, natural colorants, amino acid derivative, submerged cultivation, *in vitro* synthesis

Background

Natural food colorants are at the focal point of interest due to the growing consumer awareness of possible harmful effects of synthetic colorants [1,2]. With the increasing importance of diet and health, the food additive industry faces new challenges in providing natural colour alternatives. Currently, most industrially applied natural colorants are extracted directly from natural sources *e.g.* betanins (beet root *Beta vulgaris*), carminic acid (extracted from the female insect *Dactylopius coccus* [3]) or lycopene (tomato *Solanum lycopersicum*). Their production is highly dependent on the supply of raw ingredients and are subject to seasonal variation both in regards to quantity and quality [4]. These limitations can potentially be overcome by exploring new sources for natural pigments such as microorganisms [5]. Fungi have an excellent potential in this regard, as they are known to naturally biosynthesize and excrete diverse classes of secondary metabolites including pigments within a broad range of colours [6].

One well- studied pigment producing fungal genus is *Monascus*, which has been used for many years in the manufacturing of traditional foods in Asian countries [7]. Pigments from *Monascus* are referred to as "*Monascus* pigments" and describe a mixture of chemical compounds called azaphilones including yellow, orange, and red constituents. *Monascus* pigments have been associated with production of the mycotoxin citrinin, excluding their use as food colorants in western countries [8]. *Monascus* pigment biosynthesis is highly dependent on the medium conditions, and different colour shades can be produced through nitrogen source selection [9]. The pyranoquinone bicyclic core of the azaphilones is often highly oxygenated and can react with amines by exchanging the pyran oxygen with nitrogen [10]. This exchange leads to a colour shift from yellow/orange to red. It has been demonstrated in *Monascus* that supplementing ammonium nitrate with a specific amino acid leads to the incorporation of the amino acid into the pigment core structure resulting in the formation of new pigment derivatives in the pigment cocktail [11].

Large scale *Monascus* pigment production, therefore results in a mixture of different *Monascus* pigments [12]. The disadvantage is that the exact pigment composition is difficult to control and has high batch to batch variation. For standardised and controlled production, needed to obtain regulatory approval for food applications, it would be preferable to obtain a single pigment.

Monascus pigments were first described in association with *Monascus* species, but have henceforth been identified in numerous other species [13]. One promising *Monascus*-like pigment producer was identified by *Mapari et al.* under the

name *Penicillium purpurogenum* and was later reclassified by *Frisvad et al*. as a *Talaromyces atroroseus* IBT 11181 [14,15]. When cultivated on plates, *T. atroroseus* excreted a bright red colour and in submerged cultivation *T. atroroseus* demonstrated promising pigment profile and, most importantly absence of mycotoxins [15]. *T. atroroseus* has been studied previously [15–23] and so far six *Monascus*- like pigments have been structurally presented [14–16,19,24]. Four of these pigments are illustrated in Fig 1. A novel class of pigments has recently been identified in *T. atroroseus.* They are referred to as atrorosins and are derivatives of the known pigment *Penicilium purpurogenum*- violet (PP-V) (*Isbrandt et al*. in progress).

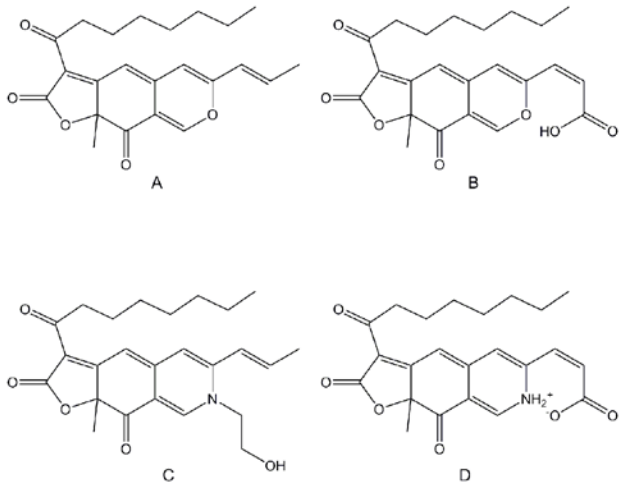


Figure 1. Chemical structures of *Monascus* pigments which have been identified in *T. atroroseus.* A) *Penicillium Purpurogenum* –Yellow (PP-Y)[18], B) *Penicillium Purpurogenum* –Orange (PP-O)[18], C) *Penicillium Purpurogenum* – Red (PP-R)[24], D) *Penicillium Purpurogenum* –Violet (PP-V)[16].

To tailor pigment production in submerged cultivations, both the physiological conditions for optimal fungal growth as well as the conditions for product bio-synthesis need to be considered [25]. Media constituents including carbon and nitrogen sources as well as process parameters such as pH have shown to significantly affect pigment production in *Monascus* [26–29]. However, optimal conditions for pigment production are not always in accordance with optimal conditions for growth. As with *Monascus* species, the pigment profile from *T. atroroseus* is strongly influenced by the choice of nitrogen source [30]. In *T. atroroseus*, ammonium nitrate together with yeast extract promotes production of violet/red pigments *Penicillium purpurogenum*- violet (PP-V) and *Penicilium purpurogenum*- red (PP-R)), but when yeast extract is used as sole nitrogen

source PP-V and PP-R production is replaced by production of *Penicillium purpurogenum- orange* (PP-O) and *Penicilium purpurogenum*- yellow (PP-Y) giving a yellow/orange colour profile [18].

The aim of this study was to investigate the effect of using amino acids as the sole nitrogen source on the pigment biosynthesis in the promising *T. atroroseus* IBT 11181. Amino acids were screened in submerged shake flask cultivations as the sole nitrogen source to assess the differences in biomass accumulation and pigment production. Based on these results, two cultivation methods to produce single atrorosins, instead of a cocktail of pigments as seen in *Monascus*, has been developed. Finally, to further investigate the atrorosin biosynthesis and mechanism of amino acid incorporation, *in vitro* synthesis of atrorosins was demonstrated using PP-O and an amino acid as precursors, yielding pure atrorosins.

Methods

Reagents

All reactants and media ingredients were purchased from Sigma-Aldrich GmbH, Steinheim, Germany.

strain and propagation

The strain used in this study was *Talaromyces atroroseus* IBT 11181 (DTU strain collection). *T. atroroseus* spores were propagated on CYA agar plates and incubated at 30 °C for 7 days. Spores were harvested with 0.9% sodium chloride solution (NaCl), filtered through mira-cloth, centrifuged and then re-suspended in 0.9% NaCl solution. The spore concentration was determined by using a Burker-Turk counting chamber. All cultivations were inoculated to give an initial spore concentration of 10^6 spores/ ml.

Submerged Cultivation

The medium for cultivation in shake flasks was composed of sucrose (7.5 g/L), glucose (0.375 g/L), KH_2PO_4 (10 g/L), NaCl (1 g/L), $MgSO_4 \cdot 7\ H_2O$ (2 g/L), KCl (0.5 g/L), $CaCl_2 \cdot H_2O$ (0.1 g/L) and trace metal solution (2 mL/L). The trace metal solution consisted of $CuSO_4 \cdot 5\ H_2O$ (0.4 g/L), $Na_2B_4O_7 \cdot 10\ H_2O$ (0.04 g/L), $FeSO_4 \cdot 7\ H_2O$ (0.8 g/L), $MnSO_4 \cdot H_2O$ (0.8 g/L), $Na_2MoO_4 \cdot 2\ H_2O$ (0.8 g/L), $ZnSO_4 \cdot 7\ H_2O$ (8 g/L). The nitrogen source was 0.1 M each of L-amino acid. 0.1 M of KNO_3 was used as the control. The pH of the medium was adjusted to pH 5 with aqueous NaOH and HCl. Cultivations were carried out in non-baffled shake flasks at 30 °C and 150 rpm in rotary shaking incubators (Forma orbital shaker, Thermo Fisher Scientific, US) with a volume of 100 ml. Samples were taken after 96 hrs. Shake flask experiments were carried out in triplicates.

Submerged cultivation with serine as the sole nitrogen source was carried out in a 1 L bioreactor. For cultivation in bioreactors, 20 g/L sucrose was used, but the other constituents were kept at the concentrations used in shake flasks. The cultivation was run at 30 °C, 800 RPM, 1 VVM and a pH of 4.5. The bioreactor experiments were carried out in duplicates.

Two-step cultivations were carried out in 1 L bioreactors. The medium for two-step cultivation contained sucrose (20 g/L), glucose (1 g/L), KH_2PO_4 (10 g/L), NaCl (1 g/L), $MgSO_4 \cdot 7\ H_2O$ (2 g/L), KCl (0.5 g/L), $CaCl_2 \cdot H2O$ (0.1 g/L) and trace metal solution (2 mL/L). For the first step, 2 g/L of KNO_3 were used as nitrogen source. After 53 h of cultivation the concentration of PP-O in the medium was measured and a 10-fold concentration of serine was added (1 g/L) to induce formation of the amino acid derivative and to ensure optimal conversion. The cultivation was performed at 30 °C, 800 RPM, 1 VMM and a pH of 4.5.

All reactor based cultivations were carried out in duplicates using Sartorius 1 L bioreactors (Satorius, Stedim Biotech, Goettingen, Germany) with equivalent working volumes and equipped with 2 Rushton six-blade disc turbines. The pH electrode (Mettler Toledo, OH/USA) was calibrated according to manufacturer's standard procedures. The bioreactor was sparged with sterile atmospheric air and off-gas concentrations of oxygen and carbon dioxide were measured with a Prima Pro Process Mass Spectrometer (Thermo-Fischer Scientific, Waltham, MA/USA), calibrated monthly with gas mixtures containing 5 % (v/v) CO_2, 0.04 % (v/v) ethanol and methanol, 1 % (v/v) argon, 5 % (v/v) and 15 % (v/v) oxygen all with nitrogen as carrier gas (Linde Gas, AGA, Enköping, Sweden). The pH was controlled by automatic addition of 2 M NaOH and H_2SO_4.

Sampling

Samples for dry weight (DW) analysis, HPLC, absorbance analysis and LC-MS analysis were taken regularly throughout the cultivations. Samples intended for HPLC, absorbance and LC-MS were filtered through a sterile Statorius Stedim filter (BRAND) with a pore size of 0.45 µm in order to separate biomass from the filtrate.

Dry Weight analysis

DW was assessed on filters which were pre-dried in a microwave for 20 min, kept in a desiccator for a minimum of 10 min and weighed. For DW analysis, the filters were placed in a vacuum filtration pump and ca. 10 ml of culture broth was added. Subsequently the filters with the biomass were dried in a microwave for 20 min and kept in a desiccator for a minimum of 10 min before being re-weighed. The weight of the biomass was determined as the difference of the filter weight before and after sample application.

Analysis of extracellular metabolites by HPLC

Culture samples were filtered through a 0.45 µm cellulose acetate filter (Frisenette, Knebel, Denmark). The samples were frozen and kept at -20°C until analysis. Glucose, glycerol, pyruvate, succinate, acetate and ethanol were detected and quantified using an Agilent 1100 HPLC system equipped with a refractive index and Diode array detector (Agilent Technologies, Waldbronn, Germany) and with an

Aminex HPX-87H cation-exchange column (BioRad, Hercules, Ca, USA). Compounds were separated by isocratic elution at 30°C, with 5 mM H2SO4 at a flow rate of 0.8 mL min-1. Quantification was performed using a six-level external calibration curve with glucose and pyruvate detected at a wavelength of 210 nm and sucrose, fructose, succinate, glycerol, acetate and ethanol by refractive index measurements.

Extraction and purification of Pigments

Filtered and centrifuged fermentation broth was extracted three times, with 1/3 volume of EtOAc, at pH 3 adjusted with formic acid (FA). The combined EtOAc phases were evaporated to 100 mL and extracted twice with milli-Q water (1:1) at pH 8 adjusted with ammonium hydroxide (NaOH). The water phase was readjusted to pH 3 with FA and extracted two times with EtOAc, followed by evaporation, to yield >95% pure pigment (a mixture of atrorosins and N-amino acid monascorubramine, ratio>10:1). The two pigments were separated on a Gilson 332 semi-prep HPLC system equipped with a Gilson 172 diode array detector, using a LUNA II C18 column (250 mm x 10 mm, 5 µm, Phenomenex), with a water/acetonitrile gradient.

LC-MS analyses

Ultra-high Performance Liquid Chromatography-High Resolution Mass Spectrometry (UHPLC-HRMS) was performed on an Agilent Infinity 1290 UHPLC system (Agilent Technologies, Santa Clara, CA, USA) equipped with a diode array detector. Separation was obtained on an Agilent Poroshell 120 phenyl-hexyl column (2.1 × 250 mm, 2.7 µm) with a linear gradient consisting of water and acetonitrile both buffered with 20 mM FA, starting at 10% B and increased to 100% in 15 min where it was held for 2 min, returned to 10% in 0.1 min and remaining for 3 min (0.35 mL/min, 60 °C). An injection volume of 1 µL was used. MS detection was performed in positive detection mode on an Agilent 6545 QTOF MS equipped with Agilent Dual Jet Stream electrospray ion source with a drying gas temperature of 250 °C, gas flow of 8 L/min, sheath gas temperature of 300 °C and flow of 12 L/min. Capillary voltage was set to 4000 V and nozzle voltage to 500 V. Mass spectra were recorded at 10, 20 and 40 eV as centroid data for *m/z* 85–1700 in MS mode and *m/z* 30–1700 in MS/MS mode, with an acquisition rate of 10 spectra/s. Lock mass solution in 70:30 methanol:water was infused in the second sprayer using an extra LC pump at a flow of 15 µL/min using a 1:100 splitter. The solution contained 1 µM tributylamine (Sigma-Aldrich) and 10 µM Hexakis (2,2,3,3-tetrafluoropropoxy)phosphazene (Apollo Scientific Ltd., Cheshire, UK) as lock masses. The $[M + H]^+$ ions (*m/z* 186.2216 and 922.0098 respectively) of both compounds was used.

Quantitative Analysis of the Pigment

The absorbance values of the individual pigment solutions were determined using a Synergy 2 photo spectrum (BioTek, Germany) and a 96 well microtiter plate. 150 µL of sample broth of each amino-acid-pigment-solution were scanned in the range of 200-700 nm and maximum absorbance values were determined. Absorbance at 500 nm indicated presence of red pigments. A standard curve of an orange

and red pigment was used to calculate the concentration in the medium. For the amino acids, where no standard curve was available the absorbance is given in AU/150µL.

Chemical derivatization of PP-O to form atrorosins

300 µL of a 0.05 M solution of different amino-group containing reactants were added to 300 µL of a 0.003 M solution of PP-O. The tested reactants were L-glutamine (pH 2 and pH 9), L- glutamic Acid (pH 2 and pH 9), L- ornithine (pH 2 and pH 9), L- tryptophan (pH 2 and pH 9), L- serine (pH 2), and D- serine (pH 2). The pH was adjusted to pH 9 with formic acid. The samples were vortexed and analysed by LC-MS.

Results

Influence of amino acid supplements on biomass accumulation and pigment production

To assess the potential of the canonical amino acids as sole nitrogen source for biomass and pigment production, all 20 natural amino acids and ornithine were tested in shake flask experiments in duplicates. A control containing KNO_3 as nitrogen source was also cultivated to benchmark biomass and pigment production. Accumulated biomass values and the absorbance intensities for all tested 21 amino acids and the control are shown in Fig 2.

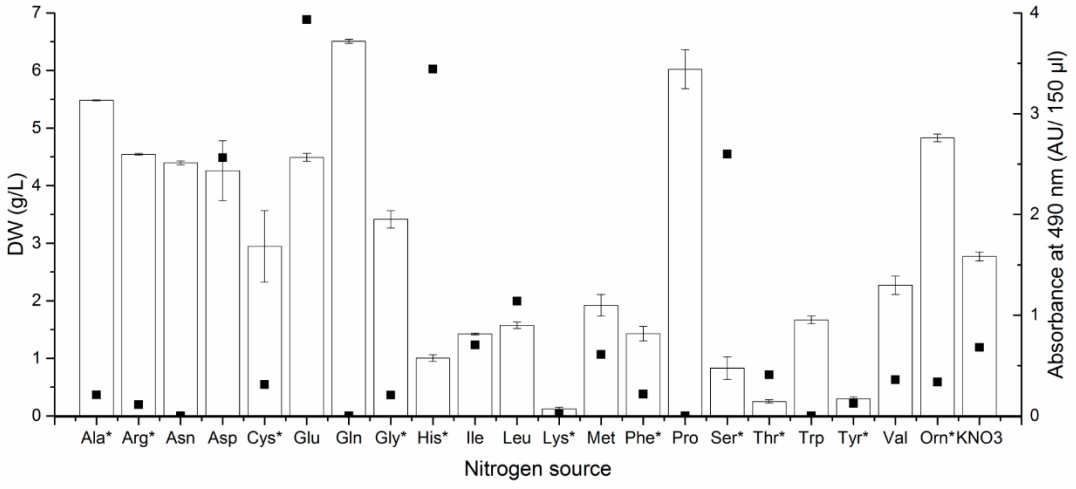


Figure 2. Maximum Biomass accumulation (bar diagram) after 96 hrs and absorbance intensity of fermentation broth (▪) of *T. atroroseus* cultures in Shake Flasks with single amino Acids as sole nitrogen source. * indicates absorbance data for 120 hrs is plotted, where the maximum value was obtained.

Biomass accumulation was highest with glutamine (6.55 ± 0.3 g/L), followed by proline (6.05± 0.3 g/L), alanine (5.48 ± 0.01 g/L) and ornithine (4.83 ± 0.07 g/L). The use of arginine (4.45 ± 0.02 g/L), asparagine (4.4 ± 0.03 g/L), aspartic acid (4.26 ± 0.5 g/L) and glutamic acid (4.49 +0.24 g/L) gave similar results. The control with KNO_3 as nitrogen source yielded only 2.77 +0.06 g/L biomass. Histidine, isoleucine, leucine, lysine, methionine, phenylalanine, serine, threonine, tryptophan and tyrosine had low biomass accumulation below 2 g/L.

Regarding pigment production, amino acids which yielded high biomass concentrations did not necessarily produce high amounts of pigment. There was no pigment production with proline and tryptophan, and very low absorbance intensities were measured when adding alanine and ornithine. Other low absorbance intensity (below 1) yielding amino acids were arginine, asparagine, cysteine, glutamic acid, glycine, isoleucine, lysine, methionine, phenylalanine, tyrosine and valine. The control with potassium nitrate had an absorbance intensity of 0.68 AU/150µl. The highest absorbance intensity was obtained with glutamic acid (3.9 AU/150µl), followed by aspartic acid, histidine, and serine. In contrast to glutamic acid, glutamine did not lead to high pigment concentration.

Biomass accumulation peaked at around 96 hrs, thereafter a stationary growth phase and decline was observed. For some amino acids, pigment production was ongoing throughout the stationary phase and peaked at around 120 hrs. This indicated that the biosynthesis of the pigments continued after the exponential growth phase. Data for absorbance intensity/ biomass are shown in table 1 for the 4 amino acids with the most favourable ratios (histidine, serine, threonine, glutamic acid and the control KNO_3; other data not shown).

TABLE 1 ABSORBANCE INTENSITY/ BIOMASS, NAME AND [M/Z] OF THE CORRESPONDING ATROROSIN OF THE 4 HIGHEST YIELDING NITROGEN SOURCES AND THE CONTROL

Nitrogen source	Yield Y_{xp} (Intensity (AU/150µl) / biomass g/L)	Name of atrorosin(s) detected	[m/z] (Da)
His	3.43	Atrorosin- H	m/z= 550.2184
Ser	3.13	Atrorosin- S	m/z= 500.1915
Thr	1.63	Atrorosin- T	m/z= 514.2072
Glu	0.88	Atrorosin- E	m/z= 542.2021
KNO_3	0.25	PP-V	m/z= 412.1755
		Atrorosin- T	m/z= 514.2072
		Atrorosin- Q	m/z= 541.2181
		Atrorosin- S	m/z= 500.1915
		Atrorosin- E	m/z= 542.2021

LC-MS ANALYSIS OF FERMENTATION BROTH FROM SHAKE FLASKS

When *T. atroroseus* was cultivated with potassium nitrate as nitrogen source, a plethora of different azaphilone pigments was produced (Fig. 3 A). This cocktail of pigments contained not only the novel group of pigments called atrorosins but also known monascorubramine derivatives. The most abundant pigments from *T. atroroseus* on potassium nitrate were PP-V and atrorosin derivatives from threonine, glutamine, serine and glutamate. Apart from glutamine, these amino acids were the same high yielding amino acids from the one step cultivation presented in table 1. Atrorosin- H (histidine derivative) was, however, not detected in the pigment mixture. On the contrary, when a single amino acid was used as the sole nitrogen source, only one single type of azaphilone pigment was produced, namely the respective atrorosin with the fed amino acid. UV-VIS- chromatograms at 520 nm from the cultivations on histidine, serine, threonine, glutamate and KNO_3 are shown in Fig. 3 A). In these chromatograms compounds that absorb light at 520 nm and therefore appear red were detected. As only one or two peaks were present in the chromatograms for the single amino acid cultivation, the purity of the fermentation broth with regard to azaphilone was greatly increased compared to the KNO3 control. The highest peak corresponded to the produced atrorosin. The smaller peak in the cultivation with histidine, serine and threonine corresponded to an amino-acid

derivative without the carboxylic acid. All atrorosins, regardless of their amino acid incorporation, had absorbance maxima at 245 nm, 275nm, 420 nm and 520 nm. As an example, the absorbance spectrum from atrorosin-E is shown in Fig. 3 B).

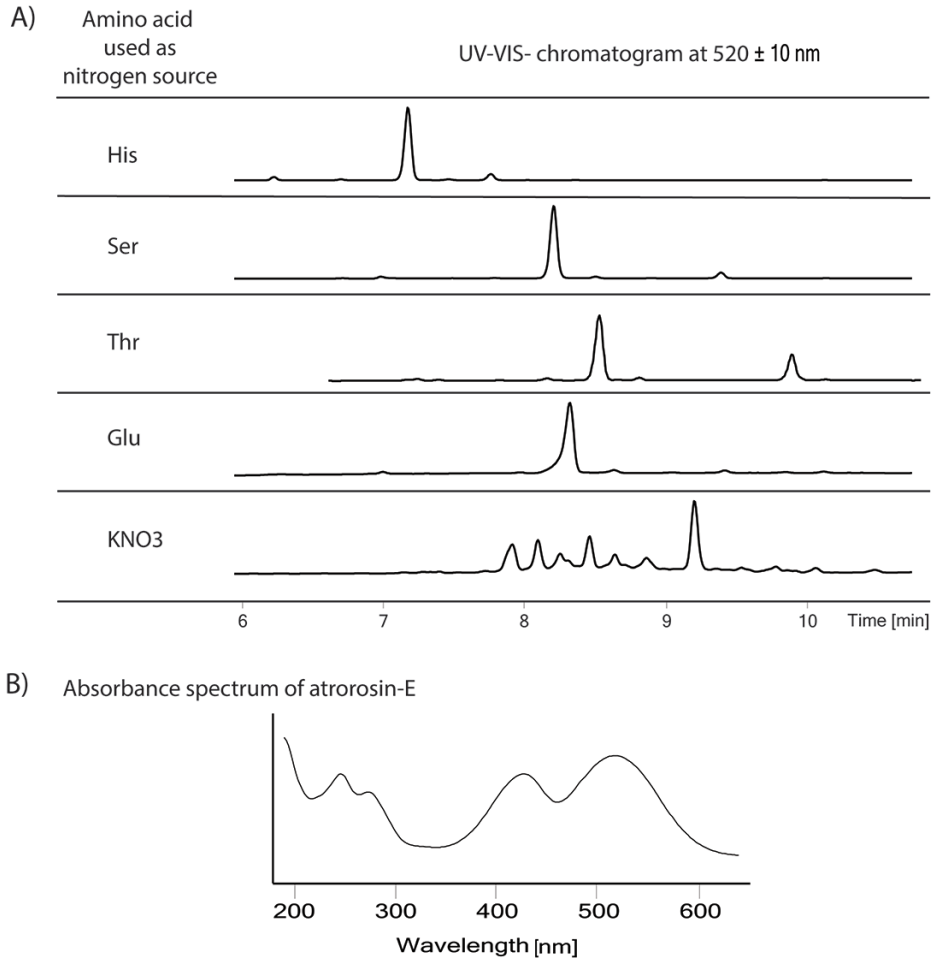


Figure 3. A) UV-VIS-Chromatograms (at 520±10 nm) of fermentation broth from cultivation with KNO_3 or amino acids as nitrogen source. B) Absorbance spectrum of Atrorosin-E.

LAB-SCALE AMINO ACID-DERIVATIVE PRODUCTION IN THE BIOREACTOR

The superior pigment purity with serine as the sole nitrogen source was the basis for bioreactor cultivations. Cultivation in bioreactors allows for pH control, controlled aeration and sparging which should lead to both higher biomass and pigment production. Two cultivation methods were developed, one using only serine as the nitrogen source similar to the shake flask and secondly a two-step process with an initial concentration of potassium nitrate to allow for biomass accumulation before a nitrogen switch to serine. Previous experiments indicated that under nitrogen limiting conditions, synthesis of PP-O was

favoured and no nitrogen was incorporated in the pigment core structure. This knowledge was applied in the two-step cultivation. First, PP-O production was promoted in the first step by inducing nitrogen-limited conditions and then, in a second step, with the addition of serine, PP-O was converted into atrorosin-S. Cultivation with KNO_3 as nitrogen source served as a control.

Fig. 4 shows the cultivation and pigment profiles. The cultivation shown in Fig. 4 A) had 0.1 M KNO_3 as nitrogen source, the cultivation profile as shown in Fig. 4 B) used 0.1 M serine as sole nitrogen source and cultivation shown in Fig. 4 C) started with an initial concentration of 0.02 M KNO_3 and after depletion approx. 55 hours, 0.01 M of sterile serine was added to the bioreactor.

The cultivation with KNO_3 as sole nitrogen source reached carbon depletion approx. 75h, however CO_2 production continued until 120h before a sudden drop in CO_2 was observed. DW accumulation and pigment production reached its maximum after 130h, after CO_2 production had dropped. The cultivation yielded around 5 g/L of DW and 0.35 g/L of red pigment mixture. During the cultivation pigment changed colour from orange (PP-O) to red (atrorosins) as carbon was depleted as seen in Fig. 4 D). Just as observed in the shake flask cultivation, a mixture of red pigments was produced (Fig. 4 G).

Cultivation with serine as the sole nitrogen source reached carbon depletion at 180 h, yielding 0.9 g/L of atrorosin-S and 6.5 g/L of biomass (Fig. 4B). While cultivation duration was much longer compared to KNO_3, most of this can be attributed to a very long lag phase of almost 80 hours. Atrorosin-S production increased with fungal growth during the entire time course of the fermentation and no PP-O was observed as seen in Fig. 4E), and it is noteworthy to state that there was minimal atrorosin impurity as seen in Fig. 4 H).

Initial conditions of the two-step cultivation process included low amounts of potassium nitrate (2 g/L), which allowed the fungus to propagate. In this process, the start of the cultivation is similar to the control with a short lag phase of 30 hours before growth. The cultivation propagated for 55 hours producing the orange pigment PP-O shown in Fig. 4I). As indicated from the CO_2 trace a limitation was reached at around 55 hours, and growth entered stationary phase. Based on the low amounts of potassium nitrate compared to the control, it was assumed that the nitrogen was depleted. At this time point, serine was added to the cultivation at a concentration of 1 g/L serving as new sole nitrogen source. Growth of the fungus continued and serine was incorporated into the PP-O azaphilone core structure, making atrorosin-S as seen in Fig. 4F & 4I).

At the end of the cultivation, all PP-O was converted into atrorosin-S as can be seen on the UV-chromatogram in Fig. 4I). The two-step cultivation yielded 0.9 g/L of atrorosin-S, and 7.4 g/L biomass. Both, biomass accumulation and pigment production peaked at the time of carbon depletion at approx. 100h.

During the first phase of the cultivation on KNO_3 two isomers of PP-O were produced seen in Fig. 4I). After the cultivation was switched to serine as nitrogen source, both isomers of PP-O were converted into only one isomer of atrorosin-S. This suggested a stereospecific conversion of PP-O to atrorosin-S by an enzyme. The peak at 8 minutes in the chromatogram corresponded to atrorosin-S confirmed by mass spectrometry ((*Isbrandt et al*. in progress) and the two peaks at 11 minutes corresponded to two isomers of PP-O confirmed by mass spectrometry ((*Isbrandt et al*. in progress).

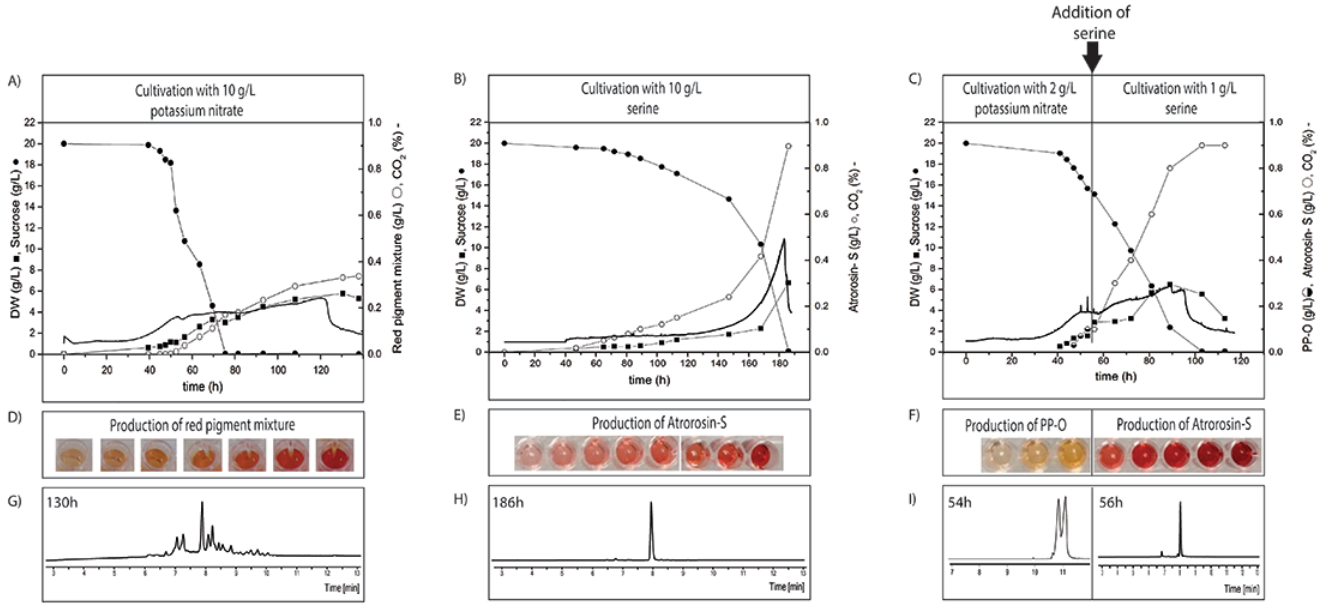


Figure 4. A) Time course of cultivation with KNO_3. B) Time course of cultivation with serine C) Time course of cultivation with nitrogen source shift (two-step cultivation) D) Colour profile of fermentation supernatant from cultivation with KNO_3 E) Colour profile of fermentation supernatant cultivation with serine F)) Colour profile of two-step cultivation G) UV chromatogram (520±10nm) from fermentation supernatant from cultivation with KNO_3 at the end of the cultivation H) UV chromatogram (520±10nm) from fermentation supernatant from cultivation with serine at the end of the cultivation I) UV chromatogram (520±10nm) from fermentation supernatant from two-step cultivation before and after the addition of serine

CARBON TO PIGMENT CONVERSION

Industrially it is relevant to evaluate the performance of processes based on the level of carbon converted to pigment. Carbon to pigment ratio was calculated for the three bioreactor cultivations performed along with additional cultivation data from NH_4NO_3 and $(NH_4)_2SO_4$ from previous cultivations at pH 5 in 1L bioreactors. Table 2 summarizes carbon conversion and evaluates pigment purity . The media composition was the same except that 0.1 M of NH_4NO_3 and $(NH_4)_2SO_4$ was used as respective nitrogen sources.

TABLE 2 PERCENTAGE OF CARBON CONVERTED TO PIGMENT ON THE DIFFERENT NITROGEN SOURCES

Nitrogen Source	KNO_3	NH_4NO_3	$(NH_4)_2SO_4$	One-step cultivation Serine	Two-step cultivation KNO_3 &Serine
Carbon conversion to pigment	1.2 % Pigment mixture	0.5% Pigment mixture	0.8% Pigment mixture	5% Atrororosin-S	6% Atrororosin-S

Cultivations with inorganic nitrogen sources (potassium nitrate, ammonium sulfate, and ammonium nitrate) all produced a mixture of pigments. The total yield of pigment was estimated using a standard curve of atrorosin-S for quantification. It is clear that introduction of amino acids either as the initial nitrogen source or as added at depletion of inorganic nitrogen source gives much higher yields almost 4-fold higher, while at the same time yielding highly pure pigment profiles.

IN VITRO SYNTHESIS OF ATROROSINS

Atrorosins as amino acid derivatives of PP-O was demonstrated to be chemically prepared by addition of PP-O with an amino containing molecule under basic conditions. Derivatisation of oxygen containing *Monascus* pigments by reaction with nitrogen containing molecules such as amino acids, amino sugars or other primary amines has previously been speculated [31]. In this study we successfully demonstrate synthesis of six atrorosins by addition of PP-with amino acids in a basic environment pH 9 as seen in Fig. 5). The amino acids selected for synthesis were L-glutamate, L-glutamine, L-ornithine, L-tryptophan, L-serine and D-serine. Fig. 5 shows UV chromatograms (520 nm± 10) of the 6 atrorosins synthesized *in vitro*, as well as a control containing only the precursor PP-O.

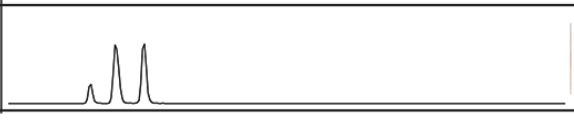
Amino Acid precursor	M (Da)	synthesized Atrorosin	UV- VIS- Chromatogram at pH 9	Color after reaction
	413.1595	control PP-O		
L-Serine	500.1915	Atrorosin-S		
D-Serine	500.1915	Atrorosin-S		
L-Glutamate	542.2021	Atrorosin-E		
L-Glutamine	541.2189	Atrorosin-Q		
L- Tryptophan	599.2388	Atrorosin-W		
L- Ornithin	527.2388	Atrorosin-Orn		

Figure 5. UV-VIS- chromatograms (520 nm ± 10) and corresponding masses of the 6 synthesized analogues and the precursor PP-O at basic conditions. Elemental composition calculated from LC-HRMS (data not shown).

In vitro synthesis was not affected by the L- or D-form of the amino acid, as both enantiomers from serine were successfully incorporated and both resulted in two isomers of atrorosin-S. Glutamine and tryptophan, which did not lead to high atrorosin production in shake flask experiments, could however be used as precursors for atrorosin synthesis *in vitro*. Similarly, these results demonstrate that the non-proteinogenic amino acid ornithine could be incorporated *in vitro*. PP-O was under basic conditions instantly converted fully to the corresponding atrorosin, with the exception of glutamine where a small amount remained. Both isomers of PP-O were equally well converted and resulted in two isomers of the atrorosin. The two isomers of tryptophan co-eluted. Ornithine, which contains two primary amines, led to 4 isomers, of which two were also co-eluting.

Test for presence of Citrinin and other mycotoxins

Other azaphilone producers such as various *Monascus* species are known to also produce the mycotoxin citrinin, making them unsuitable for use in European and American food products. Citrinin was not detected in any of the pigment extracts

from *T. atroroseus*. Fig. 6 shows extracted ion chromatograms (EIC's) of citrinin (m/z = 251.0290) of the one- and two-step production procedures with serine, as well as of the control process using potassium nitrate as the nitrogen source. It clearly demonstrates that citrinin is not produced under any of the three cultivation conditions.

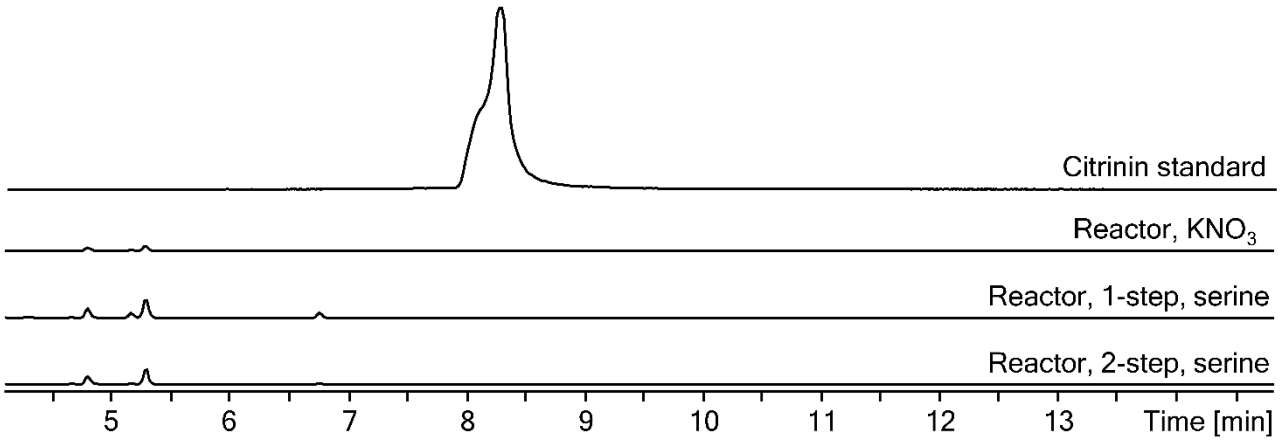


Figure 6. Extracted ion chromatograms of citrinin (m/z = 251.0290) for an authentic standard versus the one- and two-step cultivation, as well as a for a standard procedure using potassium nitrate as the nitrogen source.

DISCUSSION

The aim of this study was to investigate the effects of nitrogen sources on pigment biosynthesis in submerged cultivation of *T.atroroseus* and to design a process where specific pure pigments could be produced. Finally the mechanism of incorporation of amino acid into the core structure of atrorosins was investigated.

INFLUENCE OF NITROGEN SOURCE ON PIGMENT BIOSYNTHESIS

Submerged cultivation of *T. atroroseus* in shake flasks with different amino acids as sole nitrogen source, clearly shows that biomass accumulation and pigment production are not correlated (Fig. 2). While all but two amino acids could be used for propagation of the fungus but only five amino acids promoted high amounts pigment. Individual use of proline, lysine, asparagine and tryptophan as the sole nitrogen source, did not result in pigment biosynthesis. We speculate that proline could not be incorporated into the pigment core structure because it is a secondary amine, as primary amines have been reported necessary for incorporation.

The differences in pigment yields across the amino acids could be affected by the dynamic environment in shake flasks which have no pH control, suboptimal mixing and limited oxygen saturation throughout cultivation. The final pH of the cultures varied greatly, likely due to the different nature of amino acids and the way they are metabolized and their charge (Data not shown). As reported in previous studies [32–34], it was already hypothesized from the shake flask studies in this work, that pH is an important parameter for pigment biosynthesis but the optimal pH for pigment production does not necessarily correspond to the optimal pH for fungal growth [35]. To assure reproducible pigment production on different nitrogen sources, cultivation of *T.atroroseus* in controlled bioreactors with pH control was essential.

Analysis of the produced pigments, regardless of the amino acid as nitrogen source, demonstrates that cultivations with individual nitrogen sources produced their respective amino acid derivatives of PP-O. Interestingly, absorbance spectra of all amino acid derivatives were identical, with absorbance optimum of 510 nm. This is in contrast to Jung *et al.* 2003, who reported that amino acids like serine and glutamine would promote red colour while phenylalanine, isoleucine, leucine and valine would promote orange and yellow colour hues[11]. We suspect that the orange colour of their supernatant was due to incomplete incorporation of the respective amino acids, and were in fact due to the presence of PP-O, monascorubramine, and other orange and yellow pigments in their samples. In their experiments, Jung *et al.* used a *Monascus* sp., which was cultivated with both, amino acids and low amounts of ammonium nitrate. Depletion of ammonium nitrate was not reported, suggesting that the orange colour was caused by ammonium and not by the amino acid derivative. Several studies demonstrated that the use of ammonium nitrate resulted in the formation of orange pigments by *Monascus* sp [36]. By cultivation with amino acids as sole nitrogen source we have successfully avoided this phenomenon of cocktail pigments yielding pure atrorosins which all have a red chromophore.

Process protocol for high purity yielding azaphilone production

Building on the ability to produce pure atrorosins in shake flasks, two experimental procedures were hypothesized and tested in 1L bioreactors. First, a one-step cultivation using only serine as nitrogen source was tested, and secondly a two-step process, where an initial amount of potassium nitrate was used for biomass formation was followed by addition of serine to induce production of atrorosin-S. Both of these two novel methods successfully tailor the production to make pure atrorosin-S. Both, the one- and the two-step cultivation methods were conducted at a pH of 4.5 and resulted in significantly improved yields in regard to purity and quantity of atrorosin- S compared to other inorganic nitrogen sources. Comparison

to control cultivations with potassium nitrate as nitrogen source, HPLC analysis demonstrate a mixture of different red atrorosins and PP-V (table 1), whereas cultivations with serine as sole nitrogen source or induced at potassium nitrate depletion, only atrorosin-S was detected in the fermentation broth. All previously reported cultivation protocols in the literature on both, *Monascus* or *Talaromyces*, resulted in the production of a pigment cocktail probably incorporating any molecules containing amino-groups. When providing only one amino-containing molecule in excess, this was preferred in the incorporation of PP-O.

MECHANISM OF AMINO ACID INCORPORATION

It is highly intriguing that only a subset of amino acids was incorporated into the core of PP-O when cultivated in shake flaks. Besides proline not having a primary amine, we expected amino acids such as asparagine, glutamine, tryptophan and lysine to be incorporated into the core structure. To address this, the precursor PP-O was purified for *in vitro* synthesis of atrorosins by addition of amino acids.

Glutamate, L-serine, and ornithine were used as positive controls, whereas glutamine, tryptophan, and D-serine were tested to evaluate incorporation. PP-O was reacted with the respective amino acid in a basic aqueous solution at pH 9. As demonstrated in Fig. 5, addition of amino acids successfully incorporated into PP-O producing the respective atrorosins. It has been reported in other studies that *in vitro* conversion from orange into red *Monascus* pigments can be done [11,37,38], however *in vitro* synthesis of PP-V by addition of PP-O with glutamine in water has so far not been successfully conducted [21]. Xiong *et al.* report that a pH above 4 is critical for incorporation of amino acid both *in vitro* and *in vivo* [37]. Our results support that transamination is favoured at pH conditions above 4 and *in vitro* synthesis was successfully demonstrated at pH 9.

The results of this study are the first step in understanding the mechanism of amino acid incorporation into PP-O. PP-O as produced in the bioreactor exists as two isomers but subsequent addition of amino acid to the cultivation as seen in bioreactor converts PP-O to the respective atrorosin but only in one isomer form. Similar, in all one-step cultivation, both in bioreactors and shake flasks, only one atrorosin isomer was detected. These results suggest an enzymatic conversion of PP-O into the specific atrorosins under presence of the fungus. This hypothesis is in accordance with the literature [21]. However, *in vitro* incorporation of amino acids in PP-O into atrorosins, the two isomers of PP-O were converted into two entiomers of the respective atrorosins at pH 9, with the exception of tryptophan. The resulting two isomers of atrorosins *in*

vitro and the only one isomer *in vivo*, further suggests that *in vivo* incorporation of amino acid into PP-O is enzymatic. We speculate that amino acids which were not incorporated or minimally incorporated in the shake flask experiment as seen in Fig. 2, is due to poor metabolisation of the respective amino acids leaving little free amino acid available for incorporation into atrorosins, or the respective amino acids alter the environment significantly, such as pH, as not to allow for precursor production.

CONCLUDING REMARKS

Talaromyces atroroseus is a promising candidate for the production of novel natural pigments for industrial applications, as it excretes large amounts of water-soluble pigment into the medium. Current cultivation protocols with *T. atroroseus* use potassium nitrate, ammonium sulphate, or complex nitrogen sources and result in a mixture of different azaphilone pigments. We have in this study successfully demonstrated that by using a single amino acid as the sole nitrogen source, only one type of pigment is produced, namely the atrorosin amino acid derivative. This process allows for the desired amino acids derivative to be tailor-made by choice of amino acid. The high degree of purity of the atrorosin in the fermentation broth greatly eases downstream processing and purification, which are a necessity if this is to be industrially relevant for the food ingredient industry [39].

Another important matter for the production of azaphilone pigments is safety, and citrinin has been widely associated with *Monascus* pigments. As seen in this study *T. atroroseus* does not produce citrinin under the cultivation processes described here, which greatly benefits *T. atroroseus* as an industrial fungal cell factory for pigment production. Production of pure products has priority for approaching regulatory path for approval of novel natural colorants for foods. In this study, two cultivation methods for production of a class of new compounds, atrorosins, in high purity was presented. For industrial relevance, the two-step cultivation method is to be preferred as it is cheaper in terms of nitrogen source, because only 2 g/L of amino acid compared was used compared to 10g/L from the one-step cultivation. Furthermore, the two-step cultivation is shorter by 80 h. To further increase the industrial potential, a pre-culture could be implemented into the process potentially further reducing cultivation time [11,19,20].

List of abbreviations:

PP-O	*Penicillium purpurogenum*- orange
PP-V	*Penicillium purpurogenum*- violet
PP-R	*Penicillium purpurogenum*- red
PP-Y	*Penicillium purpurogenum*- yellow
LCMS	Liquid chromatography mass spectrometry
DW	Dry weight

Declarations

Authors contributions

GT, AO, MW initiated and coordinated the etinre project. All authors revised the manuscript. GT, AO, TI, TOL, MW conceived and designed the experiments. GT, AO and TI performed the experiments. GT, AO, TI, TOL, MW analyzed the data. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

All data and materials are available from the corresponding authors on reasonable request.

Consent for publications

All co-authors have agreed to the content and form of the manuscript for publication.

Ethics approval and consent to participate

This article does not contain any studies with human participants or animals performed by any of the authors.

Funding

Funding were from the Technical University of Denmark Department of Bioengeering (DTU Bioengineering).

References

1. Amchova P, Kotolova H, Ruda-Kucerova J. Health safety issues of synthetic food colorants. Regul. Toxicol. Pharmacol. Elsevier Ltd; 2015;73:914–22.

2. Scotter MJ. Emerging and persistent issues with artificial food colours: natural colour additives as alternatives to synthetic colours in food and drink. Qual. Assur. Saf. Crop. Foods. 2011;3:28–39.

3. Dapson RW. The history, chemistry and modes of action of carmine and related dyes. Biotech. Histochem. 2007;82:173–87.

4. Wissgott U, Bortlik K. Prospects for new natural food colorants. Trends Food Sci. Technol. 1996;7:298–302.

5. Dufossé L. Microbial production of food grade pigments. Food Technol. Biotechnol. 2006;44:313–21.

6. Dufossé L, Fouillaud M, Caro Y, Mapari SA, Sutthiwong N. Filamentous fungi are large-scale producers of pigments and colorants for the food industry. Curr. Opin. Biotechnol. 2014;26C:56–61.

7. Arunachalam C, Narmadhapriya D. Monascus fermented rice and its beneficial aspects: A new review. Asian J. Pharm. Clin. Res. 2011;4:29–31.

8. Liu BH, Wu TS, Su MC, Ping Chung C, Yu FY. Evaluation of citrinin occurrence and cytotoxicity in Monascus fermentation products. J. Agric. Food Chem. 2005;53:170–5.

9. Shi K, Song D, Chen G, Pistolozzi M, Wu Z, Quan L. Controlling composition and color characteristics of Monascus pigments by pH and nitrogen sources in submerged fermentation. J. Biosci. Bioeng. Elsevier Ltd; 2015;120:145–54.

10. Stadler M, Anke H, Dekermendjian K, Reiss R, Sterner O, Witt R. Novel bioactive

azaphilones from fruit bodies and mycelial cultures of the ascomycete Bulgaria inquinans (Fr). Nat. Prod. Lett. 1995;7:7–14.

11. Jung H, Kim C, Kim K, Shin CS. Color characteristics of monascus pigments derived by fermentation with various amino acids. J. Agric. Food Chem. 2003;51:1302–6.

12. Woo PCY, Lam C-W, Tam EWT, Lee K-C, Yung KKY, Leung CKF, et al. The biosynthetic pathway for a thousand-year-old natural food colorant and citrinin in Penicillium marneffei. Sci. Rep. 2014;4:6728.

13. Gao J, Yang S, Qin J. Azaphilones: Chemistry and Biology. Chem. Rev. 2013;113:4755–811.

14. Frisvad JC, Yilmaz N, Thrane U, Rasmussen KB, Houbraken J, Samson R a. Talaromyces atroroseus, a New Species Efficiently Producing Industrially Relevant Red Pigments. PLoS One. 2013;8:e84102.

15. Mapari SA, Meyer AS, Thrane U, Frisvad JC. Identification of potentially safe promising fungal cell factories for the production of polyketide natural food colorants using chemotaxonomic rationale. Microb. Cell Fact. 2009;8:24.

16. Ogihara J, Fujimotoz Y. Biosynthesis of PP-V, a Monascorubramine Homologue by Penicillium sp. AZ. J. Biosci. Bioeng. 2000;90:2–4.

17. Ogihara J, Kato J, Oishi K, Fujimoto Y. PP-R, 7-(2-hydroxyethyl)-monascorubramine, a red pigment produced in the mycelia of Penicillium sp. AZ. J. Biosci. Bioeng. 2001;91:44–7.

18. Ogihara J, Oishi K. Effect of ammonium nitrate on the production of PP-V and monascorubrin homologues by Penicillium sp. AZ. J. Biosci. Bioeng. 2002;93:54–9.

19. Ogihara J, Kato J, Oishi K, Fujimoto Y, Eguchi T. Production and structural analysis of PP-V, a homologue of monascorubramine, produced by a new isolate of Penicillium sp. J. Biosci. Bioeng. 2000;90:549–54.

20. Arai T, Umemura S, Ota T, Ogihara J, Kato J, Kasumi T. Effects of Inorganic Nitrogen Sources on the Production of PP-V [(10Z)-12-carboxyl-monascorubramine] and the Expression of the Nitrate Assimilation Gene Cluster by Penicillium sp. AZ. Biosci. Biotechnol. Biochem. 2012;76:120–4.

21. Kojima R, Arai T, Matsufuji H, Kasumi T, Watanabe T, Ogihara J. The relationship between the violet pigment PP-V production and intracellular ammonium level in Penicillium purpurogenum. AMB Express. 2016;6:43.

22. Arai T, Kojima R, Motegi Y, Kato J, Kasumi T, Ogihara J. PP-O and PP-V, Monascus pigment homologues, production, and phylogenetic analysis in Penicillium purpurogenum. Fungal Biol. Elsevier Ltd; 2015;119:1226–36.

23. Arai T, Koganei K, Umemura S, Kojima R, Kato J, Kasumi T, et al. Importance of the ammonia assimilation by Penicillium purpurogenum in amino derivative Monascus pigment, PP-V, production. AMB Express. AMB Express; 2013;3:19.

24. Ogihara J, Kato J, Oishi K, Fujimoto Y. PP-R, 7-(2-hydroxyethyl)-monascorubramine, a red pigment produced in the mycelia of Penicillium sp. AZ. J. Biosci. Bioeng. 2001;91:44–7.

25. Papagianni M. Fungal morphology and metabolite production in submerged mycelial processes. Biotechnol. Adv. 2004;22:189–259.

26. Said FM, Brooks J, Chisti Y. Optimal C:N ratio for the production of red pigments by Monascus ruber. World J. Microbiol. Biotechnol. 2014;30:2471–9.

27. Lin TF, Demain AL. Effect of nutrition of Monascus sp. on formation of red pigments. Appl Microbiol Biotechnol. 1991;36:70–5.

28. Broder CU, Koehler PE. Pigments produced by Monascus purpureus with regard to quality and quantity. J. Food Sci. 1980;45:567–9.

29. Yoshimura M. Production of Monascus-Pigment in a Submerged Cultures. Agr.Biol.Chem.

1975;39:1789–95.

30. Santos-Ebinuma VC, Roberto IC, Simas Teixeira MF, Pessoa A. Improving of red colorants production by a new Penicillium purpurogenum strain in submerged culture and the effect of different parameters in their stability. Biotechnol. Prog. 2013;29:778–85.

31. Lin TF, Yakushijin K, Bt GH, Demain AL. Formation of water-soluble Monascus red pigments by biological and semi-synthetic processes. 1992;9:173–9.

32. Carels M, Shepherd D. The effect of pH and amino acids on conidiation and pigment production of Monascus major ATCC 16362 and Monascus rubiginosus ATCC 16367 in submerged shaken culture. Can. J. Microbiol. 1978;24:1346–57.

33. Carels M. The effect of changes in pH on phosphate and potassium uptake by Monascus rubiginosus ATCC 16367 in submerged shaken culture. 1979;25.

34. Méndez A, Pérez C, Montañéz JC, Martínez G, Aguilar CN. Red pigment production by Penicillium purpurogenum GH2 is influenced by pH and temperature. J. Zhejiang Univ. Sci. B. 2011;12:961–8.

35. Santos-Ebinuma VC, Roberto IC, Francisca M, Teixeira S, Jr AP, Bioquímico-farmacêutica DDT, et al. Improvement of submerged culture conditions to produce colorants by Penicillium purpurogenum. Brazilian J. Microbiol. 2014;742:731–42.

36. Lin TF, Demain AL. Negative effect of ammonium nitrate as nitrogen source on the production of water-soluble red pigments by Monascus sp. Appl Microbiol Biotechnol. 1995;43:701–5.

37. Xiong X, Zhang X, Wu Z, Wang Z. Accumulation of yellow Monascus pigments by extractive fermentation in nonionic surfactant micelle aqueous solution. Appl. Microbiol. Biotechnol. 2014;99:1173–80.

38. Jang H, Choe D, Shin CS. Novel derivatives of monascus pigment having a high CETP

inhibitory activity. Nat. Prod. Res. 2014;28:1427–31.

39. Tolborg G, Isbrandt T, Larsen TO, Workman M. Establishing novel cell factories producing natural pigments in Europe. In: Singh O V., editor. Biopigmentation Biotechnol. Implement. Wiley, NJ, USA; 2017. p. 23–60.

Appendix 3

ESTABLISHING NOVEL CELL FACTORIES PRODUCING NATURAL PIGMENTS IN EUROPE

Gerit Tolborg, Thomas Isbrandt Petersen, Thomas Ostenfeld Larsen, Mhairi Workman

TABLE OF CONTENTS

1. INTRODUCTION

One of the most distinctive features of manufactured food is its colorful visual perception, thus the purchase behavior of modern consumers is driven frequently by appearance. In the supermarket, a particular color can be associated with freshness and quality of the product. But as consumer awareness increases regarding the link between diet and health, the eventual harmful effects of synthetic colorants are increasingly problematic. The food additive industry thus faces new challenges in providing natural color alternatives and the replacement of chemically synthesized dyes with bio-derived ones.

The market research company marketsandmarkets estimated the global natural food color market to be worth 1.1 billion US$ in 2014 and predicted to reach 1.7 billion US$ by 2020[1]. In 2014 the market of natural colorants drew level with the market share of synthetic colorants as both represented 34% of the overall food color market (Figure 1). The key applications of food colors are confectionary and bakery, beverages, packaged foods, dairy products and others such as frozen foods, condiments and dressings, functional food and pet food. Beverages are the second largest category occupying around 20% of the total market share[1].

The controversial topic of synthetic dyes in food has been the subject of debate for many years[2] and it seems that natural or nature-identical colorants have a more healthy image in the eye of the consumer[3]. A recent nationally representative "Consumer Reports" survey in the US of 1005 adults found that 62% of consumers usually seek out products with a "natural" food label[4].

It is clear that consumer demand for natural food colorants and transparent labelling will be major drivers for facilitating growth of the natural food color sector.

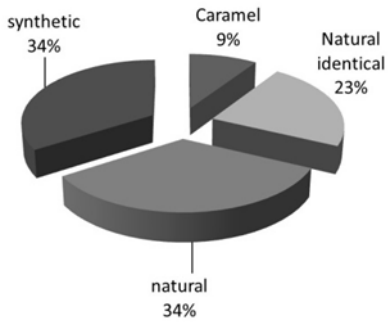


FIGURE 1. PERCENTAGE MARKET SHARE OF FOOD COLORANTS IN 2014[5]

Currently, most natural colorants are extracted from fruit skins, roots or seeds, which make their production dependent on the supply of raw ingredients which can have large variations in quantity and quality[6]. These limitations could be overcome by employing the techniques of industrial biotechnology to meet the production demand for quantity and variety of natural pigments as food colorants. With this approach suitable, robust microbial hosts (as "cell-factories") would be cultivated in industrial scale fermenters in order to satisfy the demand for consistency in quality and color hues and generate versatile pigments tailored for the food industry.

Fungi are known to naturally biosynthesize and excrete diverse classes of secondary metabolites (SMs) including pigments with an extraordinary range of colors and chemical classes. Among the SMs produced are carotenoids, melanins, flavins and quinones[7]. Fungi, particularly ascomycetes and basidiomycetes, are known to biosynthesize and secrete SMs with bioactivity giving them an advantage in natural ecological niches[8]. So far, this potential has remained more or less untapped as a source of pigments in the food industry. Furthermore, production costs of these pigments should be competitive with those of synthetic pigments or those extracted from a natural source. The successful use of fermentation physiology together with metabolic engineering could allow the efficient production of pigments synthesized by filamentous fungi[9]. So far, there are only limited reports on fungal pigment producing bioprocesses which are known to operate on an industrial scale[10]. One example is the application of species within the fungal genus *Monascus*.

Solid state fermentation of rice by *Monascus* has a long tradition in East Asian countries which dates back at least to the first century A.D. *Monascus* fermented rice products are called “ang-kak” and they are used as a food colorants for yellow, orange and red color hues[11]. Monascus pigment manufacturing companies include Tianyi Biotech, Shandong Zhonghui, Wuhan Jiacheng, Henan Zhongda, Kiriya Chemical and Yiyuan Food Chemical[12]. In 2007, Tiyani Biotech reported an annual productivity of 1500 tons of powdered *Monascus* red rice employing both solid-state and submerged fermentaion[13]. *Monascus* pigment production is however associated with the harmful mycotoxin citrinin and as a result, these pigments are not approved for human consumption in Europe and the USA[14].

Blakeslea trispora was presented as a new natural source of *β*-carotene in 1995[7] and has been reported to produce up to 44.5 mg per g biomass[15,16]. The fungus is used in submerged fermentation processes by the company DSM as a cell factory for food colorants[14,17], however most of carotenoids on the market are still produced by chemical synthesis. Furthermore, the Czech company Ascolor described a process for the production of an anthraquinone type molecule using *Penicillium oxalicum* in 2004, which was called “Arpink Red”[18]. However, this process is no longer in production.

Clearly, the implementation of new fungal microbial cell factories for safe and reliable color production is important for the advancement of safe, biobased alternatives to chemically synthesized pigments for the European market. There is a strong interest in systematically investigating novel pigment producing cell factories. The challenge here will be to define standardized and reproducible submerged cultivation processes which will support obtaining the detailed quantitative data on physiological properties needed to both understand and improve novel fungal cell factories.

This chapter will evaluate the potential of new red pigment producers and review the current status for production of biobased pigments focusing on those resembling the *Monascus* derived pigments used in Asia. Several candidates which already have been found will be presented and main achievements will be summarized and put in context relative to the well-studied *Monascus* species.

2. COLORANTS

2.1 CLASSIFICATION OF COLORANTS

Both pigments and dyes are colorants. These terms are, however, often used interchangeably for substances responsible for coloration of the medium within which they are applied. The difference between the two is that dyes are soluble in the applied medium and pigments are not. Hence a colorant can change from being a dye to being a pigment. For example

carotenoids are dyes in oil but pigments in water. For biological pigments this distinction is normally not used and all colorants are referred to as pigments.

Pigments are compounds capable of absorbing visual light and thereby changing the color of the reflected light. The reflected light is then observed by the human eye as the color of the light not absorbed *e.g.* a pigment absorbing light in the low wavelengths of the visual light (480-540 nm) will appear red.

The ability for pigments to absorb light and thereby appear colorful is linked to the chemical structure of the compound, more specifically to the system of conjugated double bonds, known as the chromophore. When double bonds appear in a conjugated system, electrons can delocalize across the system of overlapping *p*-orbitals. When irradiated by light, the electrons can absorb photons of specific wavelengths depending on the size of the conjugated system. As a result, only some wavelengths are reflected leading to the colored appearance of the molecule.

Food colorants can be categorized as natural, nature-identical and synthetic colorants. Natural colorants are pigments that are found in nature because they are biosynthesized by a living organism. Naturally derived colorants are mainly plant extracts or pigments from plants, *e.g.* red from paprika or beetroots; yellow from saffron; orange from annatto; green from leafy vegetables[9].

Nature-identical colorants are chemically synthesized pigments, with identical chemical structures to colorants found in nature. Examples include *β*-carotene and riboflavin. This group of colorants also includes pigments resulting from chemical modifications of natural colorants. Synthetic colorants are purely man-made and do not occur in nature. Synthetic colors such as Red 40 (Allura red AC), Red 3 (Erythrosine), Blue 1 (Brilliant blue FCF), Blue 2 (Indigotine), Green 3 (Fast green FCF), Yellow 6 (Sunset yellow FCF) and Yellow 5 (Tartrazine) are used widely in the industry as coloring agents for in cosmetics, drugs, candies, beverages and many other foods.

Colorants to be applied in the Western food industry need to comply with standards and regulations. Legislation ensures that only certain pigments are permitted as food colorants and compound-specific purity standards are subject to strict regulation. Furthermore quality control is an important factor. This is mainly related to color strength, hue and intensity where certain criteria also have to be met. As the pigments are produced by microorganisms, the proof of absence of pathogens and mycotoxins must be rigorously verified. Food pigments are additionally required to be tasteless, odorless and unreactive with other constituents of the food. They should be stable over a wide range of pH, temperature and not sensitive to light or oxygen[19].

2.2 Monascus Pigments

The term "*Monascus* pigments" refers to a mixture of azaphilones including yellow, orange, and red constituents. This group of pigments has first been described in association with *Monascus* species, but nowadays numerous other species have been linked with their production (see Table 1)[20]. Azaphilones are characterized by their pyranoquinone bicyclic core, often highly oxygenated. They are known for their ability to react with amines by exchanging the pyran oxygen with nitrogen[21]. This exchange leads to a color shift from yellow/orange to red. Up until 1973, six *Monascus* pigment compounds including two yellow ones, monascin [22] and ankaflavin [23]; two orange ones, rubropunctatin [24] and monascorubrin [25,26]; and two red ones, rubropunctamin and monascorubramin[26] were identified (Figure 2). But additionally to these six major azaphilone pigments, more than 50 related pigments exist[27,28,20]. Well- studied examples include N-glutarylmonascorubramine and N-glucosylrubropunctamin[29]. Other pigments such as the two furanoisophthalides xanthomonasin A and B[30,31] or industrially useful polyketides, such as cholesterol-lowering compounds referred to as monacolins, are also produced by this genus[32].

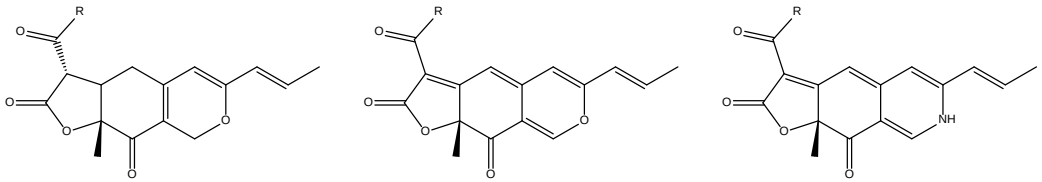


FIGURE 2. CHEMICAL STRUCTURE OF THE ORIGINAL SIX MONASCUS PIGMENTS

Many studies have been conducted on *Monascus* pigments in regard to their structure as well as to their biosynthetic pathway, optimized production strategies, detection methods,[27,33] and their biological activity[28]. *Monascus* pigment biosynthesis is considered to generally follow a polyketide pathway, but the exact mechanisms are still unclear.[27] Species used for *Monascus* pigment production include *M. pilosus*, *M. purpureus*, *M. ruber*, and *M. anka* [27].

Annual consumption of *Monascus* pigments in Japan has increased significantly over the last 30 years and new food applications like the coloration of processed meats (sausage, ham), marine products like fish paste and tomato ketchup have been reported[7]. However despite the enormous economic potential of *Monascus* pigments, they have not yet found their way into the European market, due to their association with the nephrotoxic metabolite citrinin[34]. It has been shown, that several species of *Monascus* produce citrinin[35]. Citrinin production seems to be related to pigment production and can be influenced by both media and cultivation conditions [36–38]. In order to use *Monascus* species for food colorant production for the European market, non-citrinin producing species or culture conditions unfavorable to citrinin production need to be discovered. So far, several attempts have been successful in defining the conditions for citrinin-free *Monascus* pigment production[39,40]. However, once a microorganism has been associated with mycotoxin production, the quest for FDA or EFSA approval can be a complex and elaborate process.

Instead, it seems more promising to screen for species, other than those from the *Monascus* genus, which produce *Monascus* pigments, providing a safe and viable alternative for the European market[19].

2.3 BIOSYNTHESIS OF MONASCUS PIGMENTS

Monascus azaphilone pigments are constructed from integration of two larger biosynthetic building blocks; a non-reduced polyketide and a 3-oxo fatty acid. Labeling experiments using 1-^{13}C-acetate and 2-^{13}C-acetate have shown that the biosynthetic origin of both the polyketide, , as well as of the 3-oxo fatty acid are biosynthesised from one acetyl-CoA unit, and additional malonyl-CoA units[41]. The literature on monascus pigment biosynthesis is relatively limited, and what is known about the mechanism of the pathway is therefore only speculative. However, two hypotheses predominate in the literature.

In 2014, Woo et al.[42] proposed a biosynthetic pathway after the construction of several knock-out mutants in *Penicillium marneffei*. Their work elucidated that one specific polyketide synthase (PKS) gene, namely *pks3*, was responsible for most of the biosynthesis. The proposed pathway, was based on findings by Hajjaj et al. (1999) when investigating the biosynthetic origin of citrinin in *M. ruber*[43]. As citrinin is derived from a tetraketide rather than a hexaketide, it should be noted that the pathways and modification routes might actually differ from that of monascus pigments.

The second pathway, depicted in Figure 3, is based on work done by Balakrishnan et al. in 2013 involving a gene similar to *pks3* from *P.marneffei*, namely the gene, *mpPKS5* found in *M. purpureus.*[44] The gene encodes the PKS catalyzing the formation of the backbone polyketide structure in *Monascus* azaphilone biosynthesis. Furthermore the same research group discovered a gene named *mpp7* responsible for the regioselective attachment of the 3-oxo-fatty acid.[45]

The biosynthetic pathway based on Balakrishnan et al. (2013) (illustrated in Figure 3) is proposed to start with an aldol condensation of the polyketide backbone structure, to form a cyclic structure by linking carbons 2 and 7. Next, the molecule undergoes a second condensation reaction between the oxygen atom at C-1 and C-9 to form a heterocyclic isochromene system.

During the biosynthesis several double bonds are formed by losses of water – these steps are expected to be activities of the PKS, and thus the exact order of reactions is uncertain. The oxidation at carbon 4 has been shown to be a result of tailoring enzymatic activity by the monooxygenase mppF in *M. purpureus*[46].

The final step in the biosynthesis is the attachment of the 3-oxo fatty acid to the core bicyclic pyranoquinone, expected to be synthesised by the fatty acid synthase (FAS) MpFAS2 in *M. purpureus*[46]. The fatty acid is attached to the alcohol introduced by mppF by esterification, followed by an aldol condensation assisted by the protein Mpp7 in order to control regioselectivity.[45]

The conversion from yellow and orange pigments by introduction of nitrogen is proposed to happen via a Schiff base. The conversion has been done chemically using various amino acids[47], but the biological conversion is expected to be enzyme mediated as the biological pH values are too low to favour Schiff base formation[48].

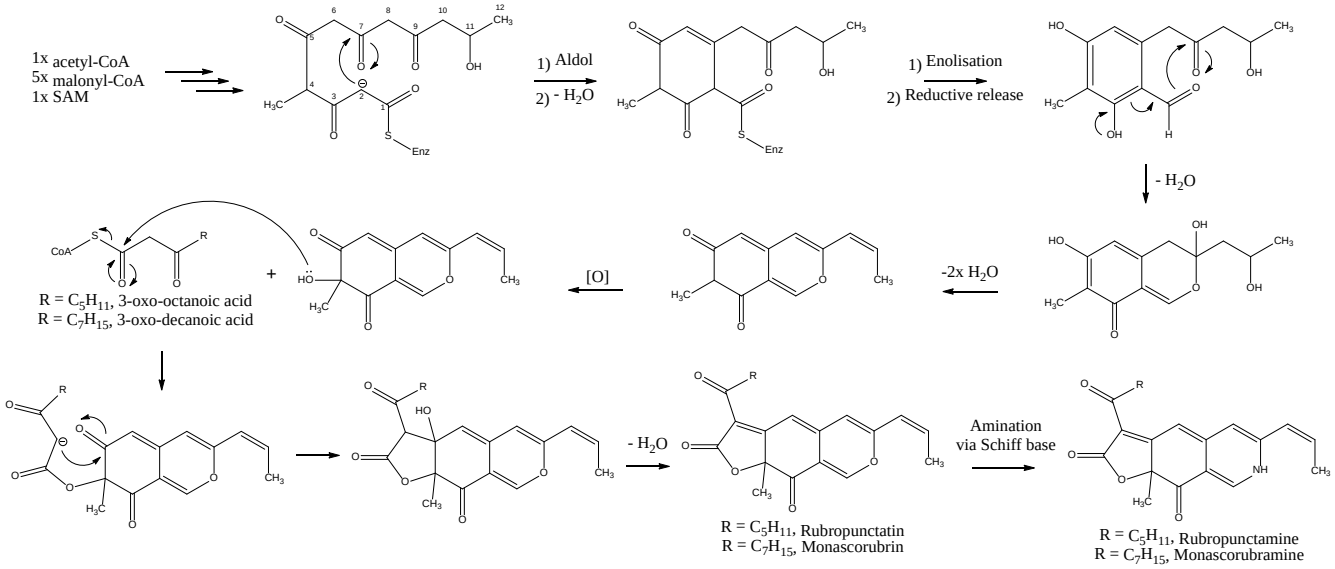


FIGURE 3. MONASCUS PIGMENT BIOSYNTHESIS BASED ON THE PROPOSAL BY BALAKRISHNAN ET AL. (2013)

2.4 DERIVATIVES OF MONASCUS PIGMENTS

Several derivatives of the six original *Monascus* pigments have been described so far. Blanc et al. (1994) reported the production of N-glutarylmonascorubramine and N-glutarylrubropunctamine by *M. purpureus* and *M. ruber* when using glutamtate as sole nitrogen source[49]. Amino acid containing derivatives originate by the substitution of the oxygen in

monascorubrine or rubropunctatine by the nitrogen of the amino group of various compounds such as aminoacids, peptides and proteins. The most widely accepted way for nitrogen incoorporation to happen is via the formation of a Schiff base in order to exchange the oxygen with nitrogen.[47,50] This substitution leads to a color change from orange to purple. Based on this principle, *Monascus* pigment derivatives containing for example glutamate[47], aspartic acid and alanine[51] have been identified and characterized. Jung et al. produced *Monascus* pigment derivatives using 20 individual amino acids in submerged cultivation as side chain precursors[52]. The pigment derivatives containing amino acids were found to be more robust towards both temperature and pH changes than the orignal pigments and showed increased photostability[53]. Moreover, various red color hues were exhibited by these pigment derivatives[47,52,53]. These properties make amino acid derivatives of *Monascus* pigments particular interesting for further exploration as food colorants. The amino acid derivatives can be obtained by either adding the amino acids into the medium[47,52] or by chemical synthesis. For the latter, the orange pigments, amino acids, and amines are dissolved in ethanol and derivatives were synthesized by mixing the orange pigment solution with either an amino acid or amine solution.[54,55]. But not only amino group-containing compounds can be used to alter the pigments. Application of a high glucose concentration to the culture medium has been shown to induce the formation of glucosyl derivatives of the pigments, such as N-glucosylrubropunctamine and N-glucosylmonascorubramine[29].

3. Screening for a *Monascus*-pigment producing cell factories for the European market

3.1 Cell factory selection and identification

In order to find a safe and promising cell factory for pigment production a chemotaxonomic screening should to be carried out. Such a screen should focus on the preselection of potential fungal pigment producers combined with the deselection of toxin producers. For many years genomic screening for pigment producers was not feasible, however with the rapidly increasing numbers of available fungal genomes, bioinformatics driven discovery of pigment producing species is now possible. The use of classical tools such as taxonomy, biochemistry and microbial physiology have so far been the methods of choice for preselecting interesting candidates[19]. These first screenings are usually performed on agar plates as they guaranteed a high throughput of various species under different media conditions. The interplay between physical and chemical parameters of a production process and the biological properties of the cell complicate cell-factory selection and promising candidates should therefore not only synthesize the product of interest but also fulfill other criteria. Desirable cell-factory characteristics include efficient substrate utilization as well as conversion of a wide substrate range, a high degree of product secretion and product stability, high yield and productivity, a minimized by-product-formation, applicability of the process in bioreactors and amenability of the organism to genetic modifications.[56]

Several species among the genus of *Penicillia* have been shown to produce *Monascus*-like azaphilone pigments without co-producing citrinin. An international patent has been filed on production of these *Monascus*-like pigments, by potentially safe strains of *Penicillium* species through the use of a combination of liquid and solid cultivation techniques[57]. Furthermore, *Talaromyces* species, e.g. *T. flavus* secretes a red pigment called mitorubrin[58]. The recently newly classified species *T. purpurogenus, T. atroroseus, T. albobiverticillius, T. minioluteus*, and *T. marneffei* have been linked in the literature previously under *Penicillium* species names[59]. This means that some species previously referred to *Penicillium purpurogenum* are now classified as either *Talaromyces atroroseus* or *Talaromyces purpureus.* Thus, the species *P. purpurogenum* no longer exists[59]. Confusion regarding species name can therefore still arise, because not all *P. purpurugenum* from the literature have been reclassified yet. Nonetheless, *P. purpurogenum* IBT 11181, described by Mapari et al. and the strains used by Arai and Ogihara[41,48,60–64] have been reclassified as *T. atroroseus* by Frisvard et al. [59].

Table 1 summarizes potential pigment producing *Penicillium* species and groups them according to their chemotaxonomic classification.

TABLE 1. POTENTIAL PIGMENT PRODUCERS FOR THE EUROPEAN MARKET

Published fungal name	Reported pigment	Ref
***T.atroroseus*, classified by Frisvad et al.[59]**		
T. atroroseus	mitorubrins and several Monascus pigments	[59]
Strain reclassified as *T. atroroseus* by Frisvard et al.[59]		
P. purpurogenum IBT 11181	Monascorubramine	[65]
	N- glutarylmonascorubramine	[65]
	N- glutarylrubropunctamine	[65]
	PP-R	[66]
P. purpurogenum IAM 15392		
	PP-R (*P. purpurogenum*- red, 2-hydroxyethyl-monascorubramine)	[63] [48,60–62]
	PP-V (P*. purpurogenum*- violet, 12-carboxyl-monascorubramine)	[48]
	PP-O (*P. purpurogenum*- orange)	[48]
	PP-Y (*P. purpurogenum*- yellow) (*for structures see Figure 4*)	
***P. purporogenum* Stoll**	Purpuride	[67]
Strains potentially belonging to *T.atroroseus*, but not examined by Frisvad et al.[59]		
P. purpurogenum DPUA 1275	Yellow, orange, red*	[68–71]
P. purpurogenum GH2	Red pigment production*	[72]
Penicillium spp.	extracellular pigment-production *	[73]
Other pigment producing strains		
P. aculeatum IBT 14263	Monascorubrin (orange) Xanthomonasin A (yellow) Threonin derivative of rubropunctatin (purple red)	[74]
P. marneffei	Monascorubrin, Rubropunctatin	[42,75]
P. pinophilum IBT 13104	Monascorubrin	[74]
T. purpurogenus	N-glutarylrubropunctamine Rubropunctatin	[59]

* no chemical structure of the pigments is reported and due to the new classification the exact species is uncertain

So far, *T. atroroseus* seems to be the most promising candidate for *Monascus* pigment production, due to its capacity to secrete various different pigments in high yields and its lack of citrinin production[41,59,63–65]. Figure 4 shows the chemical

structures of monascus pigment derivatives associated with *T. atroroseus*. All strains reclassified as *T. atroroseus* will be addressed as *T.atroroseus* for the remainder of this chapter; those which have not been reclassified will be referred to using the species name used in the respective publications.

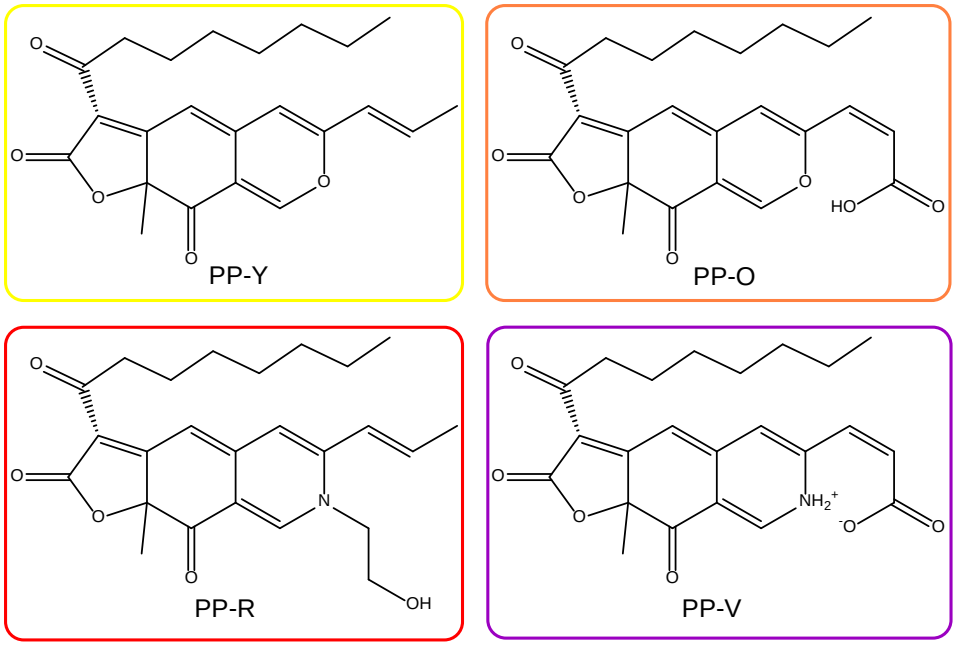


FIGURE 4. CHEMICAL STRUCTURES OF MONASCUS PIGMENT DERIVATIVES ASSOCIATED WITH T. ATROROSEUS

3.2 From single pigment producers to high performance cell factories

For more than 50 years, filamentous fungi have served as industrial cell factories and large-scale processes based on these organisms are therefore well established. One common feature of high performance cell factories is the fact that the strains have been carefully selected and were then either highly adapted or engineered to the specific conditions of the process application or the process conditions were tailored to optimize the production. The performance of filamentous fungi in lab-scale submerged cultivation determines their suitability for large-scale industrial application and is the result of complex interplay between the physical and chemical parameters of the process and the cellular biology of the fungus.

When considering pigment production from fungal hosts, several challenges exist (outlined in Figure 6). Firstly, non-domesticated species must be tested and validated as culturable in stirred tank reactors, and with natural isolates, where growth rates and morphology in submerged cultures is not known, this step is not trivial. Secondly, when color is produced by fungal species is it typically based on a number of chemical compounds, which must be separated, identified and quantified if meaningful data is to be extracted which can be further used for process and strain optimization. Preferably, single, or few, pigment compounds would be produced allowing validation and approval of bioprocesses based on one target product or product class. These challenges can be realized through integrated approaches for assessing cellular performance (quantitative physiology), genetic modification of strains (metabolic engineering), -omics analyses and modeling.

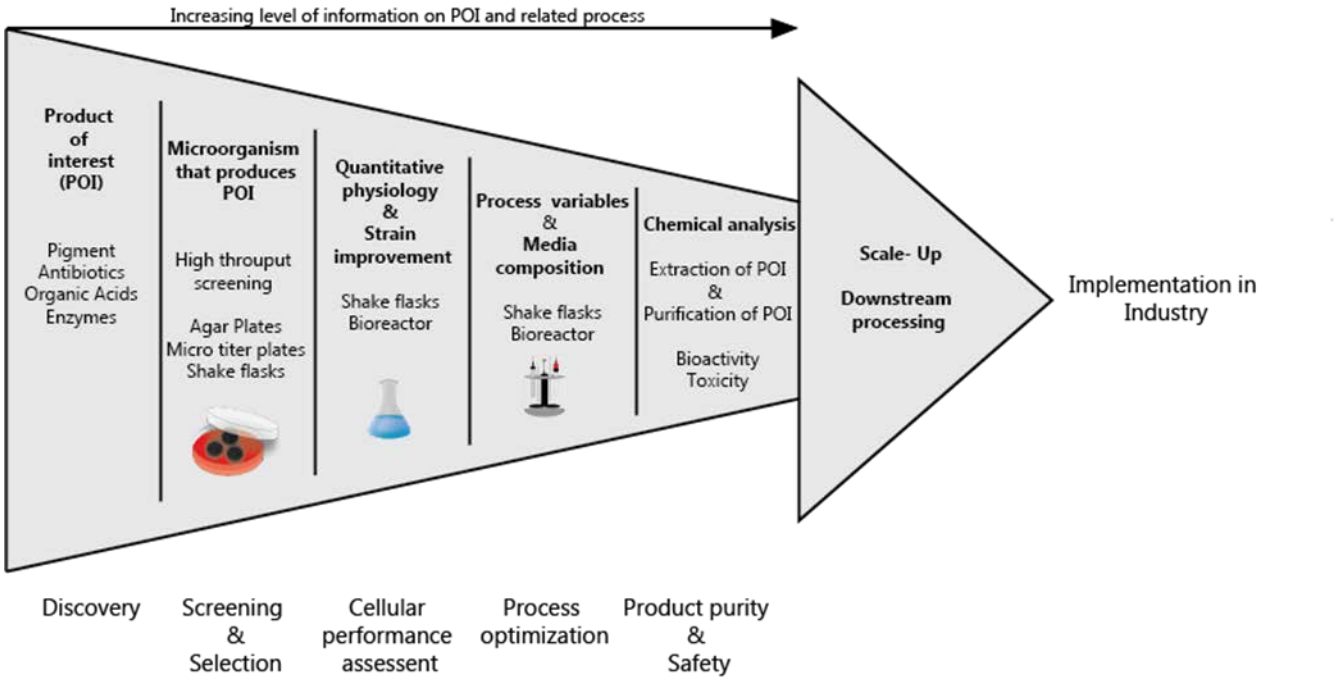


Figure 5. Major steps for establishment of a successful cell factory

Classic quantitative physiology based techniques can help designing an optimal process. The challenge is to balance cellular potential, process design, and economic feasibility[56]. So far genetic engineering approaches using *T. atroroseus* are still in their infancy, but as tools for genetic modification are constantly being developed, it is only a matter of time before direct pathway engineering is likely to be possible.

4. Assessment of the color yield

The term “monascus pigments” usually refers not only to one pure pigment but rather to a cocktail of different monascus pigments including yellow, orange and red constituents[42]. The disadvantage of mixtures is that they are difficult to control and monitor because the individual compositions might vary slightly from batch to batch. From an industrial point of view and in order to facilitate commercialization of a product it is preferable to obtain a pure pigment. Therefore some purification strategies are required to separate individual pigments from the mixture. Recovering a pure pigment allows for structural analysis and the set-up of quantification tools, such as HPLC methods. Furthermore, efficient downstream processing of the product of interest is of outmost importance for high performance cell factories. A future goal, of course, would be to design a process or organism where production of single or few pigment compounds would be guaranteed, and thus simplify validation and recovery.

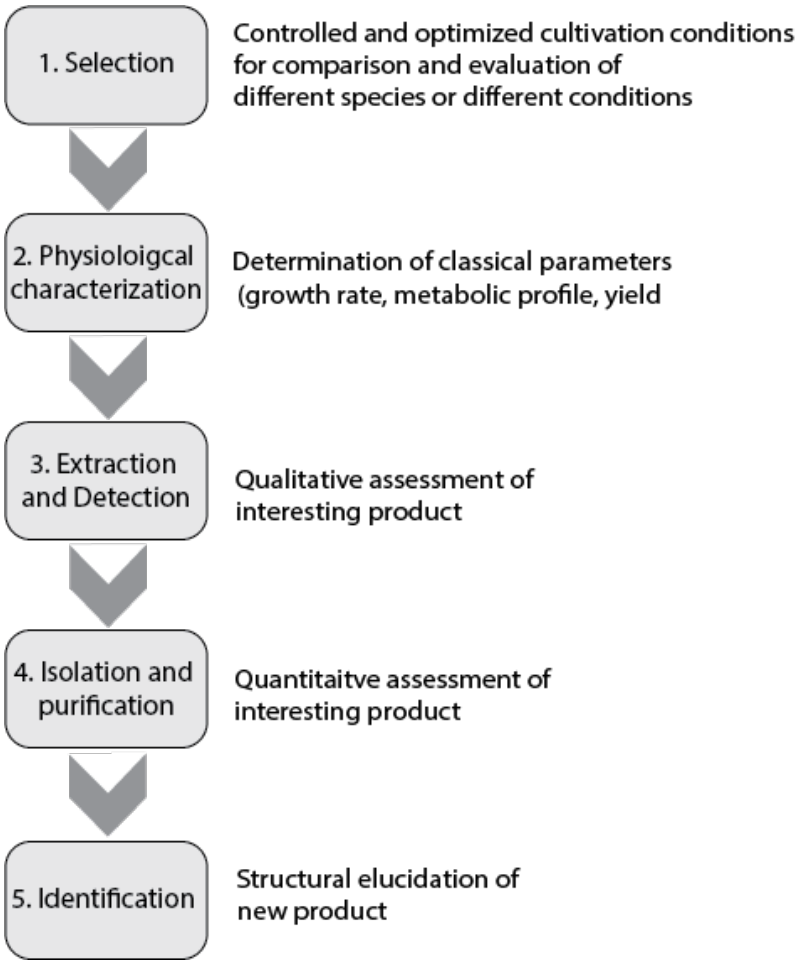


Figure 6. Steps for product identification when establishing a novel cell factory

Figure 6 here

4.1 Pigment purificationand quantification

The six original *Monascus* pigments (see Figure 2) are mostly cell bound and hydrophobic. However, they possess an aminophilic moiety that can react with amino group-containing compounds in the medium, such as proteins, amino acids, and nucleic acids to form water-soluble pigments[47]. Depending on whether the pigments are water soluble or insoluble, different extraction procedures are utilised. In order to assess the intracelullar pigments the cells need to be separated from the liquid media and extracted.[76,77] *Monascus* pigments can be extracted with methanol[78], ethyl acetate[34,49–52,55] or ethanol at various concentrations[37]. Monascus pigments from solid state fermentation and monascus pigments still bound to the mycelium from a submerged cultivation can be extracted through solid-liquid extraction[79–81] or microextraction [74,82]. Monascus pigments in the supernatant from submerged fermentation can be extracted through liquid-liquid extraction[76,79,83].

The solubility of pigments, and thereby the probability of finding them extracellularly, is generally inceased when nitrogen is incorporated, and even more so when nitrogen containing molecules, such as amino acids or glucosamines are incorporated. In contrast, pigments such as rubropunctatin and monascorubrin are less hydrophillic and therefore more often found intracellularly.

The purification of *Monascus* pigments can be done in a number of ways. The most widely applied method is using chromatography, such as thin layer chromatography which is now being gradually replaced by liquid chromatography (LC).[47,52,84,85] Additionally, capillary electrophoresis[31,86] and high-speed counter-current chromatography, have been reported as methods for analysing *Monascus* pigments.

4.1.1 Thin layer chromatography

Thin layer chromatography (TLC) is an inexpensive and widely used method for initial screening or even purification (preparative TLC) of compounds. Since the method only requires visual inspection, pigments are easily analysed. However, due to the lack of a detector, correct identification of the individual compounds is impossible. This limitation can be overcome by transferring the developed TLC spots to a suitable instrument, such as a UV-VIS spectrophotometer, mass spectrometer or NMR spectrometer. TLC can be done using both normal (e.g. silica gel) or reversed (e.g. C_{18}) phase stationary phases, depending on the desired type of separation[52,62,85]. Furthermore a variety of developing agents can be used. Different mixtures of water, chloroform, acetonitrile, methanol, tetrahydrofuran, toluene, hexane, ethyl acetate, and acetone have been reported[52,61–63,85,87].

4.1.2 Liquid chromatography

Liquid chromatography (LC) covers a wide range of techniques, from column chromatography (CC) to high performance LC (HPLC). CC is often used as an initial fractionation of crude extracts for instance using a silica gel column[64,88,89], typically followed by either TLC or HPLC[84].

For separation of more complicated mixtures of compounds, HPLC is the method of choice. As for TLC and CC both normal phase (NP) and reverse phase (RP) HPLC can be applied for analytical as well as preparative purposes. RP HPLC systems typically utilise either a water/methanol or a water/acetonitrile solvent system combined with an apolar stationary phase, such as C_{18}, C_8, phenyl-hexyl or equivalent. NP HPLC have been reported to use silica or amino columns in combination with hexane/chloroform or isopropanol/hexane, respectively.[82,85,90,91]

4.2 Detection and Identification

Ultraviolet-visible (UV-VIS) spectrophotometry is the most widely used method for detection of pigments, and by using a diode array detector (DAD), the entire absorption spectra can easily be recorded at once. Monascus pigments of different colours can be detected at different wavelengths with absorption maxima in the range between 390 and 530 nm depending on solvent and pH, representing pigments with color hues from yellow, through orange, to red/purple, as illustrated in Figure 7.[66,70,74,82]

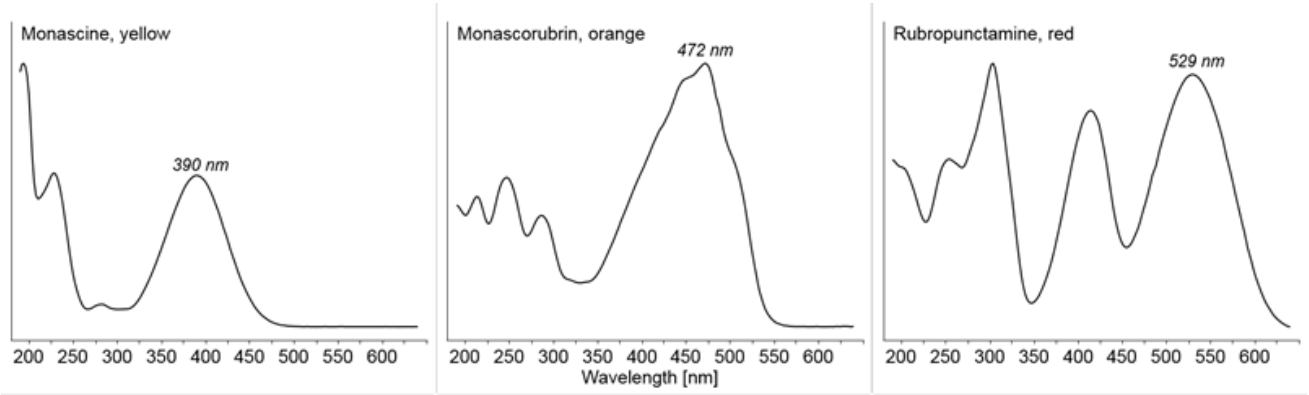


FIGURE 7. YELLOW, ORANGE AND RED UV SPECTRA FROM THREE MONASCUS PIGMENTS: MONASCINE (YELLOW), MONASCORUBRINE (ORANGE), AND RUBROPUNCTAMINE (RED)

Spectrophotometers typically give the intensity of a colour at a certain wavelenghth as absorbance (A), a unitless number that can be expressed though Lambert-Beers Law:

$$A = log(I_0/I) = \varepsilon * l * c$$

Lambert-Beers law states the the absorbance is equal to the logarithm of the ratio between I_0, the intensity of an unaffected beam of light and I, the intensity of the light after passing through the sample of interest. These values are more frequently expressed as the product of the sample concentration, c, the distance though the sample, l, and the compound specific constant ε, known as the extinction coefficitent. Since the absorption is dependent on concentration, quantification of the pigment is often based on UV-VIS measurements.

Despite the usefulness of UV-VIS detection, the method has some limitations, as several monascus pigments possess comparable absorption spectra[52]. In these cases other detection techniques, such as mass spectrometry (MS), high resolution mass spectrometry (HRMS), or infrared (IR) spectrophotometry need to be applied in order to distinguish similar compounds.

By combining several detection methods with separation strategies like HPLC, powerful means of detection can be achieved, such as HPLC-DAD-HRMS, giving information about retention time, chromophore and molecular composition simultaneously.[66,92]

4.3 QUANTIFICATION

Some of the above mentioned detection techniques can also be used for pigment quantification. A requirement for quantification of individual pigments is however their availability in pure form, otherwise only the overall quantity of pigment cocktails can be determined. Based on the UV-VIS spectrum of the pigments- in mixture or pure-, standard

curves can be used for pigment qualification using Lambert-Beers Law[93–95]. Crucial for this method is a powerful purification strategy which guarantees that only certain pigments remain in the extract. The use of MS and UV-VIS coupled to HPLC has also been reported for quantification[96]. As illustrated in Table 4.2, so far no common quantification strategy is available and the published results in terms of pigment yield are hardly comparable across research groups. In the majority of cases, the absorbance spectrum of the supernatant is detected and values measured at specific wavelengths, e.g. at 400nm, 470nm or 490nm are reported[71]. Often, there is only limited information regarding the size of the sample volume, the depth of the well or the dilution factor applied in many of the previous studies. Straight forward standard procedures for quantifying individual pigments from the fermentation broth would therefore be desirable to enable comparison of different processes across research studies.

TABLE 2. REPORTED OPTIMAL PROCESS CONDITIONS FOR *T.ATROROSEUS* AND OTHER RELATED SPECIES

Fungus	Optimal process parameters	Assessment of color	Max Color yield	Ref.
P. purpurogenum GH2	pH5, 24°C, 260h incubation time, 200rpm, 15 g/L Xylose, 3 g/L NaNO3	Absorbance measurement of supernatant (SN) at 500 nm. The yield of product per unit biomass (YP/X) was calculated as the ratio of the amount of pigment produced (Pt−P0) at a certain time (t) to the biomass generated in the same time (Xt−X0).	2.4 g/L	[72]
Penicillium spp	pH9, 30°C, 4 days incubation time, 200rpm, soluble starch, peptone	Absorbance measurement of SN at 530 nm with spectrophotometer	U_{A530}=1238	[73]
P. purpurogenum DPUA 1275	pH6.5, 30°C, 360h incubation time, 150rpm, 30g/L Sucrose, 5g/L Yeast extract, 5 mycelia discs	Absorbance measurement of SN at 400, 470, and 490 nm with spectrophotometer. Results were expressed in terms of Units of Absorbance (UA).	U_{A400}=3.08 U_{A470}=1.44 U_{A490}=2.27	[69]
P. purpurogenum DPUA 1275	pH4.5, 30°C, 336h incubation time, 150rpm, 50g/L sucrose, 10g/L Yeast extract	Same as [69]	U_{A400}=3.10 U_{A470}=2.50 U_{A490}=2.04	[71]
T. atroroseus	pH5, 30°C, 2 days incubation time, 20g soluble starch, 3g NH_4NO_3, 2g Yeast extract	Detection by thin layer chromatography, identification of PP-V and PP-R by H^1 NMR and C^{13} NMR		[48,64]
T. atroroseus	pH5, 30°C, 4 days incubation time, 200rpm, 20g soluble starch, 2g Yeast extract	Detection by thin layer chromatography, identification of PP-Y and PP-O by H^1 NMR and C^{13} NMR		[48]

T. atroroseus	Plates: pH 6.5, 25°C, 7d incubation time (CYA/YE) Shake flasks: pH 6.5, 25°C, 7d incubation time, 150 rpm (CZ)	Absorbance measurement of SN, prior adjustment to 0.40 ± 0.04 at their respective absorption maxima with purified water. Determination of CIELAB color coordinates of the same extracts using Chromameter.	(L/a*/b*/hue angle/chroma) CYA= (87.8/10.7/22.6/64.8/25.0) YE= (85.8/15.9/16.5/46.2/22.9) CZ= (85.0/19.7/22.4/46.6/29.8)	[66]

Table 4.2 here.

4.4 CIELAB

CIE L*a*b* is the name of a color space specified by the International Commission of Illumination (CIE) and it includes all perceivable colors. The coordinate L* represents the lightness of the color (L*=0, yields black and L*=100 indicates diffuse white) and a* and b* represent the color-opponent dimensions. The system is based on the fact that light reflected from any colored surface can be visually matched by an additive mixture of the three primary colors: red, green, and blue. Since the *L*a*b** model is a three-dimensional model, it can only be represented properly in a three-dimensional space. Two-dimensional depictions include chromaticity diagrams which have a fixed lightness. The CIELAB color space is used to describe fungal pigments produced under specific conditions[52,53,66,80,82] and might even give indications regarding the concentration of the pigment[97]. The extracted pigments in solution or as crude pigment extract are measured with a colorimeter to obtain L*a*b* values, which can then be used to calculate chroma (C*) and hue angle (hab) values[52]. Chroma values indicate the saturation or purity of the color. Values close to the center at the same L* value indicate dull or gray colors, whereas values near the circumference represent vivid or bright colors. This model represents a good tool to describe and assess color hue variations of pigments produced by different species or by the same species but under different culture conditions[52,53,66,80]. Its limitation is that neither structural information about the pigment, nor characteristic properties such as UV/Vis spectra about the pigment mixture nor quantitative information obtained. Its character is purely descriptive and it is therefore useful for analyzing results from initial screening experiments, but not for properly evaluating a pigment producing process.

5. Optimizing cellular performance: growth and pigment production

The optimization of cellular performance of a fungal cell factory can be tackled from many different angles. It is an interplay between media constituents, such as carbon sources, nitrogen sources and mineral salts, and culture conditions, e.g. pH, temperature and aeration rate[98–102]. Many studies on optimizing pigment production have been carried out using the one-factor-at-a-time method[69,73,103]. However this method is unable to show possible interaction effects between the selected parameters and therefore could miss promising parameter combinations. Design-of-experiment based approaches such as Plackett-Burman design[104] or fractorial factorial designs[70–72] seem more promising in finding optimal cultivation conditions, especially for novel cell factories, where little quantitative data exists in the literature.

5.1 Assessment of classical physiological parameters

A cell factory can be evaluated by key cellular performance indicators, most importantly growth rate, utilization of the substrate and yield of biosynthesized products. Different cultivation modes and conditions can be used to produce and study various physiological states. For a reliable quantitative physiological characterization, it is paramount that the cultivations are performed in a controlled environment and the reproducibility can be demonstrated. Such studies provide data on which process optimization can be based, but also give an insight into the active metabolic pathways based on the nutritional requirements and measured products[56].

Growth rate is often the key parameter used when screening potential call factories. A fast growing fungus is better suited for industrial processes because it lowers the production length and thereby production costs and the risk of contamination. Data on biomass accumulation in pigment producing *Penicillium* species or *Talaromyces* species is incomplete and so far no growth rates in bioreactors have been reported for any of the strains. Whereas pigment production in *Monascus ruber* is growth associated[95], it appears that biomass accumulation and red pigment production are not directly linked in the pigment producing *P. purpurogenum GH2*[72]. It is possible in this case that red pigment production occurs at the expense of biomass production, as in *T.* atroroseus[105]. However, this is only speculation at present as titers rather than yields have typically been reported in the literature. Mendez et al. stated 6 g/L biomass[72], Gunasekaran et al. reported 5.5 g/L biomass[73] and Santos-Ebinuma et al. reported up to 25 g/L of biomass[71]. However, these values were not related to medium composition and are thus difficult to compare. Growth rates for Monascus species are reported to vary from $0.02h^{-1}$ for *Monascus sp* [106] to 0.04^{-1} for *Monascus ruber*[29,95,107].

5.2 Media composition

Pigment synthesis is highly dependent on the medium conditions. In *Monascus,* extensive studies have been performed in order to find the optimal media conditions[95,108,109]. Some conditions favor growth, whereas others trigger pigment production. High glucose concentrations have been shown to induce formation of glucosyl derivatives of *Monascus* pigments, such as N-glucosylrubropunctamine and N-glucosylmonascorubramine[29] . Furthermore supplementing the media with a specific amino acid leads to the incorporation of the amino acids into the pigment core structure resulting in a new pigment derivative[52,103], e.g. N-glycinemonascorubramine[52] or N-glutarylrubropunctamine[49]. Media conditions of *Monascus* species serve as an excellent starting point to investigate promising carbon- and nitrogen sources for *T. atroroseus* based cell factories.

5.2.1 Carbon Source

Carbon sources for submerged cultivation of *T. atroroseus* and related species include soluble starch[64], potato starch[65], sucrose[71] glucose[64], fructose[64], galactose[64], mannose[64],arabinose[64] and xylose[72]. Gunasekaran et al. screened the effect of 11 carbon sources on pigment production and concluded that starch promoted the highest pigment yields, followed by maltose and glucose[73]. However, other studies reported that sucrose was the most promising carbon source in terms of pigment production[69]. Moreover, *P. purpurogenum* has been shown to grow on cellulose, wheat straw and wheat bran, but by using these carbon sources no pigment production was documented[110]. More systematic and quantitative studies need to be conducted to assess the influence of the carbon source on both, biomass accumulation and pigment production. Because pigment production in *Monascus* appears to be subject to strict glucose repression, which results in ethanol production if the glucose concentration in a medium exceeds 30 g/l[107], different carbon concentrations need to be tested with *T.atrororeus* to investigate this phenomenon. For *Monascus*, comparative transcriptome analysis revealed that

carbon starvation stress, resulting from the use of relatively low-quality carbon sources, contributed to the high yield of pigments by repressing central carbon metabolism and augmenting the acetyl-CoA pool[111]. Here, non-conventional carbon sources, such as ethanol may be worth investigation, as they have shown promising results for *Monascus*[112]. Since the production of secondary metabolites usually takes place in the stationary phase of growth, it may be possible to split fungal growth and secondary metabolite formation into two distinct phases. Indeed, a two-stage cultivation with maltose and ethanol was conducted successively[112]increasing efficiency of ethanol utilization for pigment production. So far sucrose seems to be the carbon source of choice for submerged processes for pigment production using *T.atroroseus*.

5.2.2 Nitrogen Source

Inorganic compounds such as ammonium chloride or ammonium nitrate and organic nitrogen like yeast extract, monosodium glutamate (MSG) and other amino acids, are good nitrogen sources for both growth and pigment production of submerged grown *Monascus spp*[27]. Furthermore, different color components (yellow, orange or red pigments) can be selectively produced through nitrogen source selection[87]. Some studies report that the presence of organic nitrogen is optimal for growth but unfavorable for pigment production. If they are used, free amino acids should be added additional to the culture medium, which then can react with the orange pigments to form red-colored complexes. However not all amino acids seem to promote pigment production equally well[113].

Reduced growth and optimized pigment formation occurs with inorganic nitrogen sources, such as ammonium chloride, sodium nitrate and ammonium nitrate[114]. Nitrates limit growth but stimulate spore and pigment formation and the use of ammonium nitrate as the nitrogen source has been found to result in the formation of mainly cell-bound orange pigments by Monascus sp[115].

Ammonium is the preferred nitrogen source over nitrate because nitrate must be reduced in an energy requiring process before it can be used for anabolic processes[107]. Use of monosodium glutamate obviates the need for its synthesis from ammonium, or nitrate via ammonium. Glutamate directly, or after conversion to glutamine, can be used for the biosynthesis of the various other metabolites and amino acids[95]. During nitrate consumption, the pH rises as nitrate is reduced to ammonium to facilitate incorporation into proteins. Therefore nitrogen sources should always be studied together with the pH[87,107,116]. Overall, monosodium glutamate seems to be the best nitrogen source for *Monascus* spp. promoting both growth and pigment production[108,115].

T. atroroseus and other related species show the same tendencies in regard to nitrogen utilization. Pigment production with *T.atroroseus* is reported on complex sources such as yeast extract[48,65] and ammonium nitrate[48,60,61,64]. Ammonium nitrate together with yeast extract promotes PP-V and PP-R production, but when yeast extract is used as sole nitrogen source PP-V and PP-R production is replaced by production of PP-O and PP-Y[48]. Ammonium and nitrate nitrogen can both be used for PP-V production, but ammonium nitrogen results in higher yields than nitrate[60].

5.2.3 Other medium components

Minerals such as magnesium sulfate, potassium chloride or phosphate can also affect pigment production with *Monascus*. Lin et al. showed that high concentrations of phosphate (above 70 mM) and magnesium sulfate (above 16 mM) have an inhibitory effect on cell growth and pigment production. In contrast, potassium chloride concentration was found not to affect cell growth or pigment production significantly[108]. The negative effects of high concentrations of phosphate and magnesium are due to inhibition of pigment synthase action[117]. The positive effects of trace metals, especially Zn^{++} have

been shown to be due to stimulation of growth and enzyme action. So far no studies on the effect of minerals on *T.atroroseus* or related species have been performed.

5.3 Cultivation parameters

In addition to the media composition, different cultivation parameters can be used to assess cellular performance. Cultivation parameters include incubation time and size and type of inoculum[69], or can involve physical parameters such as pH, temperature, oxygen supply, light or simply the cultivation mode. The reactor and impeller design can also influence product formation[81]. Many of these parameters have already been studied for *Monascus* species and from these results conclusions could be drawn to optimal cultivation conditions for *Talaromyces* related species.

5.3.1 pH

It was reported that utilization of different nitrogen sources in submerged cultivation resulted in different pH pattern affecting growth and pigment production[116]. Generally, the suitable pH for growth and pigment production of *Monascus spp.* is 5.5–6.5[27]. However, different pH values in the media may affect single *Monascus* pigment constituents, for example was ankaflavin synthesis by *M. purpureus* favored at pH 4.0, while the other pigments were independent of pH[107]. pH might also affect formation of conidia, sporulation and thereby influencing growth and pigment production[116]. By increasing the pH, the reaction between orange pigments and amino-group containing compounds to form red-colored complexes is stimulated. Conversely, a low pH prevents the nucleophilic addition of amino groups to the oxygen atoms of orange pigments and consequently red pigment formation is limited[28]. A two-stage cultivation conducted at different pH (5.5 and 8.5) resulted in increased pigment production in *Monascus purpureus*[118]. When grown in submerged conditions, *P. purpurogenum GH2* optimum for pigment production was found to be pH 5[72] and also studies with *P. Purpurogenum DPUA 1275* demonstrated that a lower pH is favorable for pigment production[71].

5.3.2 Temperature

For most microorganisms temperature is a critical environmental factor for regulating develop- mental and physiological processes*. Monascus spp*. are typically cultured at 25–30°C[119]. Nonetheless it has been reported that low temperatures (25 °C) can promote a ten-fold greater yield than higher temperatures (30°C)[120]. This could be explained by slower cell growth and improved homogeneity in the fermenter, better oxygen transfer and lower viscosity. Using solid state fermentation, optimal pigment producing conditions were reported to be 30°C indicating a thermoprotective role of the pigments[101]. *T. atroroseus, P. Purpurogenum DPUA 1275* and other related species are typically cultivated at 30°C[48,71,73].

5.3.3 Oxygen supply by aeration rate and orbital stirring

Levels of oxygen and carbon dioxide in the gas environment were found to influence pigment production significantly in *Monascus* species. Oxygen acts primarily as the final electron acceptor in oxidative phosphorylation and also as a substrate for oxygenases in fungal metabolism[121]. Hajjaj et al. reported that the concentration of biomass and secondary metabolites including monascus pigments and citrinin were increased by improving oxygen supply, especially by increasing the dissolved oxygen concentration in media through application of higher orbital stirring rates[36]. However, pigment production was increased to a lesser degree than citrinin production suggesting that a more moderate oxygen transfer coefficient would be required to improve the proportion of red pigment/citrinin production[122].

One challenge to efficient mixing and mass transfer in fungal cultivations, is the nature of the morphology in submerged cultivation. Changes in morphology can alter the viscosity of filamentous fermentation broths, with additional effects on mixing and mass transfer. Santos-Ebinuma et al. tested the significance of orbital stirring in a fractorial factorial design and concluded that while as main variable it did not have a significant effect, its interaction with other independent variables was significant[71]. Shake flask cultivation with *Talaromyces* related species are usually run at 200 rpm[72]. However, it would be recommended to conduct investigations of oxygen transfer using a bioreactor set-up because where other parameters can be controlled and dissolved oxygen concentrations as well as off-gas composition could be measured. This type of information would be essential to elucidate the relationship between oxygen, agitation and pigment production in *T. atroroseus*.

5.3.4 Light

Pigment production in *Monascus* is greatly influenced by various light sources including white, red, blue, yellow, and green light[27]. *Monascus* species generally score a maximal pigment yield in darkness and minimum one in white light[123]. *M. purpureus* and *P. purpurogenum* both yield the highest extracellular pigment yield with no light exposure[76]. Even total suppression of pigment production in direct illumination was reported[124]. It is suspected that *Monascus* species possesses a system for differential light response and regulation[125]. The responses are mediated by light photoreceptors capable of initiating the signal transmission that result in changes in the gene expression encoding enzymes responsible for mycelial growth and secondary metabolite productions in fungi[123].

Light exposure could be a relevant factor for upscaling an industrial process for pigment production in terms of production site and choice of bioreactors. So far not data is available on *Talaromyces* related species regarding light dependency.

5.4 Type of cultivation

5.4.1 Submerged vs. solid state

Monascus pigments can be produced by solid-state fermentation (SSF)[126,127] or submerged fermentation (SF)[39,103] in shake flasks or bioreactors. Whereas the products of SSF can be directly used as food colorants[128] as the fungus together with its substrate is ground to a fine powder, products of SF need to be extracted before being used.

SSF is a classical process to produce *Monascus* pigments in Asia, in which the fungus is inoculated into steamed rice or rice kernels spread on wooden trays and cultured for about 20 days in an air-, moisture-, and temperature-controlled room[7]. Other agro-industrial materials, such as potato-dextrose can be used as substrates[129]. Substrate humidity should be rather low in order to prevent bacterial contamination, circumvent the sticking together of rice grains and to keep a low glucoamylase activity[33,130]. To secure sufficient aeration of the mycelium it is advisable to separate grains from agglomerates formed during sterilization or cultivation. As the pigments produced by SSF are unpurified products their application is limited. Utilization of submerged cultivation can help to overcome the problem of process control[131].

When grown in a submerged cultivation, the fungus can be cultivated in shake flasks or bioreactors using a wide array of defined or complex media. Pigment production processes performed in bioreactors outcompete processes performed in shake flasks in terms of information level due to the higher degree of control. Improved growth and pigment production may be due to better hydrodynamics and oxygen transfer[81]. When comparing solid state with submerged fermentation,

SSF possesses many advantages including simpler technique, less capital investment, lower levels of end-product inhibition and catabolite repression, lower amount of waste output, better product recovery, and higher yield[27]. However, with respect to productivity, cultivation in a stirred tank reactor is more economically viable because it is controllable; it has shorter cultivation times, lower production costs in the long run and higher product quality. It is the only cultivation mode that assures a completely controlled process. It allows upscaling of the parameters and offers several possibilities such as continuous cultivation or fed batch cultivation[106] for studying the metabolism, fine-tuning the process and optimizing the production yield. Continuous cultivations are very attractive from an academic point of view, since the concept of steady state offers an excellent opportunity to measure the rate of metabolic reactions or characterize the morphology of the fungus at a set of well-defined operating conditions[132], however, in-depth physiological characterization would be required before continuous processes could be designed and implemented for pigment production. *T. atroroseus* and other related species have so far mainly been studied on agar plates[59,66] or in submerged cultivation using shake flasks[68,71,72] . No bioreactor data is available yet, so there are still great unexplored opportunities for studying these potential cell factories by using different submerged cultivation strategies in a bioreactor based set-ups.

5.4.2 Extractive Fermentation

Extractive fermentation offers a way to extract orange, cell bound *Monascus* pigments into the extracellular broth during cultivation by using a nonionic surfactant micelle aqueous solution[55,79,89]. A surfactant micelle aqueous solution is able to solubilize various species with a very broad polar spectrum and is so helping to overcome the reagent incompatibility- the lipophilic nature of orange *Monascus* pigments and the solubility of the hydrophilic amino acids in aqueous solution- and thereby yielding a higher red pigment formation rate.

It can be concluded that many parameters should be considered when setting up a process with a fungal cell factory (Figure 8). This requires the careful selection of potential producers followed by the identification of the product of interest. A detection, identification and purification strategy for the new product is necessary to ensure reliable quantification and qualification. The optimised process will not only be dependent on the media composition, but also on the cultivation mode and physical cultivation conditions. In order to find the best combination of these parameters statistical designs of experiment should be employed. If available, genetic engineering techniques can be used to further optimize production yields by directing the metabolism towards higher product formation or by elimination of by-products. When using a bioreactor based set-up for pigment production, investigations regarding scale-up should be undertaken in order to gain industrial attractiveness. Assessing yield coefficients of the process offers the possibility to quantitatively compare different process-designs. They can be evaluated based on product formation, substrate consumption or turn-over.

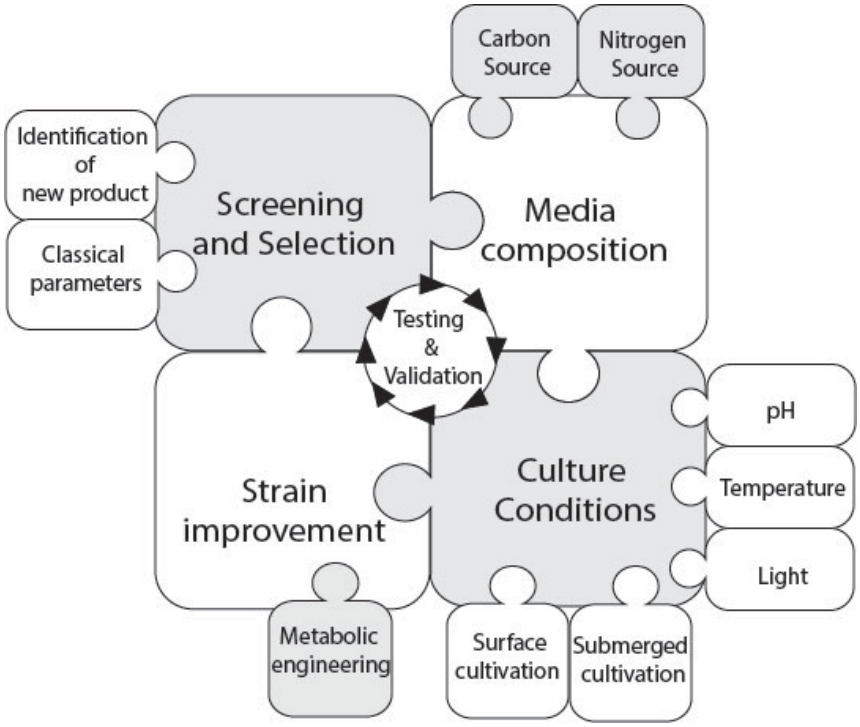


FIGURE 8. INTERACTING PARAMETERS TO BE CONSIDERED WHEN IMPLEMENTING NOVEL CELL FACTORIES FOR PIGMENT PRODUCTION

5.5 METABOLIC ENGINEERING

In the past decade, great progress has been made studying and manipulating *Monascus* spp. at the genetic level using molecular biology techniques to restrain the citrinin production or increase the yields of pigments and other beneficial products such as monacolin K. Up to now, hundreds of papers about *Monascus* molecular biology have been published[133]. For example, the polyketide syntheses genes responsible for the biosynthesis of monacolins and citrinin have been identified and cloned in *Monascus* spp.[134,135]. Furthermore, in 2013, the azaphilone pigment biosynthetic gene cluster was identified through T-DNA random mutagenesis in *Monascus purpureus* confirming a crucial role of mppR1 and MpPKS5[44]. The complete genome of *Monascus purpureus* YY-1 was provided in 2015[111]. It consists of eight chromosomes and 7491 genes. The genome from *Monascus ruber* is also publicly available[136].

Among *Talaromyces* genomes *Talaromyces stipitatus*[137,138] and *T. marneffei*[139] are publicly available, however *T. stipitatus* has never been reported to produce Monascus pigments, and *T. marneffei* is an opportunistic human pathogen. Therefore, these two are not suitable choices for *Monascus* pigment gene cluster studies. In a study in 2013, 454 sequences of *T. atroroseus* IAM15392 (Published as *P. purpurogenum*) were used as basis for qPCR expression comparison of genes (*glnA*: glutamine synthetase; *gdhA:* glutamate dehydrogenase) involved in the ammonia assimilation pathway in response to media concentration of L-glutamate and L-glutamine[61]. However, the draft sequence data has not been made publicly available. In the absence of a genome sequence for pigment producing *Talaromyces* related species, genetic engineering of these strains will be severely hampered. But even if the genomes are known, genetic tools are poorly developed for most filamentous fungi. However, accessibility of the genome of *Talaromyces atroroseus* could pave the way for omics-driven analyses which are an essential approach to evaluation of genetic regulation and offer a holistic view of cellular functions.

One potential tool for genetic engineering of filamentous fungi was presented recently. The CRISPR-Cas9 based system is now adapted for use in filamentous fungi. The system is simple and versatile, as RNA guided mutagenesis can be achieved by transforming a target fungus with a single plasmid. The system currently contains four CRISPR-Cas9 vectors, which are equipped with commonly used fungal markers allowing for selection in a broad range of fungi[140].

Gene targeting (knockout), which can be used to delete genes, replace allele, and tag genes with epitope tags or fluorescent proteins is a useful technique to study gene function in living organisms. Normally, gene replacement frequencies of filamentous fungi are very low, due to the competition with non-homologous end-joining (NHEJ)[141]. NHEJ is the dominant pathway in many eukaryotes, and its critically related genes include ku70, ku80. ku70 and ku80 proteins recruit other NHEJ proteins, therefor deletion of ku70 or/and ku80 can efficiently improve the gene replacement frequencies s of filamentous fungi[133].

Access to the genome sequence of *T. atroroseus* could greatly extend the possibilities of physiological characterization and analysis with all omics-derived technologies. Furthermore it would allow the development of metabolic engineering strategies in order to manipulate *T. atroroseus* to improve pigment secretion. Studies of the genome could also answer relevant questions, such as how many genes directly are responsible for the biosynthesis of *Monascus* pigments in *Talaromyces*, and if these are the same as in *Monascus* spp.

6. Pigment Properties

With the purpose of being used as a food colorant, *Monascus* pigments and their derivatives need to fulfill certain requirements. Heat, light and pH stability are desirable characteristics as well as good solubility in water. A potential toxicity has to be ruled out. Only few reports on toxicity of *Monascus* pigments can be found, probably due to his extensive historical use. However, the yellow pigment ankaflavin has shown selective cytotoxicity to cancer cell lines by an apoptosis-related mechanism but the structural analogue monascin has shown no cytotoxicity to all cell lines tested[142]. These results indicate however that both monascin and ankaflavin are safe for consumption in moderate concentrations. Furthermore potential bioactivity needs to be investigated and evaluated carefully in order to properly declare potential additional features of the pigment. As there are so many derivatives of the *Monascus* pigments, every modification of the molecule, meaning the incorporation of an amino acid, a carboxylic acid or a glycosyl-group might lead to slightly changed biological and chemical properties[143].

The beneficial properties of the *Monascus* product “red rice” have been demonstrated in Asia for centuries and also more recent studies demonstrate their useful biological activities for medical purposes[28,144–148]. Special health beneficial properties of *Monascus* pigments include antimicrobial, antifungal, antiviral, antioxidant, cytotoxic and anti-inflammatory activities as well as anti-mutagenic and anticancer properties or even potential anti-obesity characteristics [27,149–151]. These qualities should be further investigated in order to assess their potential impact for functional food or drug discovery.

More specific examples of pigment bioactivity are, among others, the antimicrobial effect of the two classical orange pigments rubropunctatin and monascorubrin [152,153] or of the *Monascus* pigment derivatives of hydrophobic amino acids[84]. Moreover the yellow pigments ankaflavin and monascin were found to exhibit similar anticholesterolemic effects as another secondary metabolite of *Monascus,* monacolin K[154] and experiments with rats indicated anti-obesity effects of monascus pigment amino acid derivatives with tryptophan and leucine[155].

The rather non-selective bioactivity of azaphilones is due to the formation of vinylogous γ-pyridones[149]; a reaction, in which the oxygen atom in the pyrane ring is exchanged for nitrogen from amino group-containing compounds. This explains also while when comparing the 6 classical pigments, most of the bioactive properties are associated with orange and yellow pigments, and not with red ones.

Since *T. atroroseus* and other related species produce the same group of pigments, similar properties should be associated with them. But so far, no bioactivity tests have been performed on the pigments exclusively linked to *T. atroroseus* PP-O, PP-R, PP-V and PP-Y. Unfortunately most of the pigment analysis in the literature is so far still only based on absorbance measurements, and therefore the exact composition of the pigments remains unknown. Identification of the individual pigments needs to be undertaken in order to determine detailed properties of the pigments produced.

7. Conclusion and Outlook

In this chapter the novel producer of *Monascus* pigments -*T. atroroseus*- was presented as a potential cell factory for the production of natural pigments, and a safe alternative to established processes for bioproduction of red pigments. A solid Increased knowledge of cellular performance will allow pigment production in *T.atroroseus* to be optimized and implemented on industrial scale. Firstly, standardized cultivation protocols are required to establish reproducible processes which can lead to tailored pigment production, for specific chemical compounds which can beapplied as colorants in a variety of industries. Analytical methods to analyze and confirm known molecular pigment structures and tools to elucidate novel pigments are critical for this process. Genomic data on *Talaromyces atroroseus* is necessary to fully understand the production pathway and to enable genetic engineering strain- improvement approaches. *Talaromyces atroroseus* represents an excellent candidate for satisfying the demand of society for natural derived harmless food colorants.

1. MarketsandMarkets. *Food Colors Market By Type [Natural (Anthocyanin, Carotenoid, Caramel) & Synthetic (Blue, Green, Red, Yellow)], Application (Beverage, Bakery & Confectionery, Dairy & Frozen Product, Meat Product) & Geography - Global Trends & Forecast To 2019*.
2. Kobylewski S, Jacobson MF. Food Dyes: A Rainbow of Risks. *Decis Sci*. 2010;30(2):337-360.
3. Downham A, Collins P. Colouring our foods in the last and next millennium. *Int J Food Sci Technol*. 2000;35(1):5-22. doi:10.1046/j.1365-2621.2000.00373.x.
4. Consumer Reports® national Research Center. *Natural Food Labels Survey*.; 2015.
5. Grand view Research. Food Colorants Market Analysis By Product (Synthetic, Natural, Natural Identical, Caramel), By Application (Food, CSDs, Non Alcoholic Beverages, Alcoholic Beverages) And Segment Forecasts To 2020. 2014. http://www.grandviewresearch.com/industry-analysis/food-colorants-market.
6. Wissgott U, Bortlik K. Prospects for new natural food colorants. *Trends Food Sci Technol*. 1996;7(9):298-302. doi:10.1016/0924-2244(96)20007-X.
7. Dufossé L, Galaup P, Yaron A, et al. Microorganisms and microalgae as sources of pigments for food use: a scientific oddity or an industrial reality? *Trends Food Sci Technol*. 2005;16(9):389-406. doi:10.1016/j.tifs.2005.02.006.
8. Firn RD, Jones CG. Natural products - a simple model to explain chemical diversity. *Nat Prod Rep*. 2003;20(4):382-391. doi:10.1039/b208815k.
9. Dufossé L. Microbial production of food grade pigments. *Food Technol Biotechnol*. 2006;44(3):313-321.
10. Dufossé L, Fouillaud M, Caro Y, Mapari SA, Sutthiwong N. Filamentous fungi are large-scale producers of pigments and colorants for the food industry. *Curr Opin Biotechnol*. 2014;26C:56-61. doi:10.1016/j.copbio.2013.09.007.
11. Arunachalam C, Narmadhapriya D. MONASCUS fermented rice and its beneficial aspects: A new review. *Asian J Pharm Clin Res*. 2011;4(1):29-31.
12. Reportsnreports. Organic and Monascus Pigment Industry 2014 Global Market Research Reports. http://www.prnewswire.com/news-releases/organic-and-monascus-pigment-industry-2014-global-market-research-reports-285923611.html.
13. Tianyi Biotech. Co., Ltd. http://www.made-in-china.com/showroom/tianyidg/companyinfo/Tianyi-Biotech-Co-Ltd.html.
14. Mapari SAS, Thrane U, Meyer AS. Fungal polyketide azaphilone pigments as future natural food colorants? *Trends Biotechnol*. 2010;28(6):300-307. doi:10.1016/j.tibtech.2010.03.004.
15. Lampila LEI, Wallen SE, Bullerman LB, Science F, Hall F. A review o f factors affecting biosynthesis of carotenoids by the order Monascus. *Mycopathologia*. 1985;80:65-80.
16. Berman J, Zorrilla-Lopez U, Farre G, et al. Nutritionally important carotenoids as consumer products. *Phytochem Rev*. 2014;14(5):727-743. doi:10.1007/s11101-014-9373-1.
17. Akishina RI. Blakeslea trispora producing high yield of lycopene in a suitable medium in the absence of an exogenous carotenogenesis inhibitor. 2001.
18. E. Sardaryan. Strain of the microorganism penicillium oxalicum var. armeniaca and its application. 2002.

19. Mapari S a S, Nielsen KF, Larsen TO, Frisvad JC, Meyer AS, Thrane U. Exploring fungal biodiversity for the production of water-soluble pigments as potential natural food colorants. *Curr Opin Biotechnol*. 2005;16(2):231-238. doi:10.1016/j.copbio.2005.03.004.

20. Gao J, Yang S, Qin J. Azaphilones: Chemistry and Biology. 2012.

21. Stadler M, Anke H, Dekermendjian K, Reiss R, Sterner O, Witt R. Novel bioactive azaphilones from fruit bodies and mycelial cultures of the ascomycete Bulgaria inquinans (Fr). *Nat Prod Lett*. 1995;7(1):7-14. doi:10.1080/10575639508043180.

22. Chen FC. The structure of monascin. *J Chem Soc D Chem Commun*. 1969;(3):130. doi:10.1039/c29690000130.

23. Manchard PS, Whalley WB. Isolation and Structure of ankaflavin: a new pigment from Monascus Anka. *Phytochemistry*. 1973;12(1963):2531-2532.

24. Haws EJ, Holker JSE, Kelly A, Powell ADG, Robertson A. The Chemistry of fungi.Part XXXVII. The Structure of Rubropunctatin. *J Chem Soc*. 1959;70:3598-3610.

25. Fielding BC, Haws EJ, Holker JSE, et al. Monascorubrin. *Tetrahedron Lett*. 1960;1007(5):24-27.

26. Kumasaki S, Nakanishi K. Structure of monascorubrin. *Tetrahedron*. 1962;18(5):1171-1184.

27. Feng Y, Shao Y, Chen F. Monascus pigments. *Appl Microbiol Biotechnol*. 2012;96(6):1421-1440. doi:10.1007/s00253-012-4504-3.

28. Patakova P. Monascus secondary metabolites: production and biological activity. *J Ind Microbiol Biotechnol*. 2013;40(2):169-181. doi:10.1007/s10295-012-1216-8.

29. Hajjaj H, Klaébé A, Loret MO, Tzédakis T, Goma G, Blanc PJ. Production and identification of N-glucosylrubropunctamine and N- glucosylmonascorubramine from Monascus ruber and occurrence of electron donor-acceptor complexes in these red pigments. *Appl Environ Microbiol*. 1997;63(7):2671-2678.

30. Sato K. Novel Natural Colorants from Monascus anka U-1. *Heterocycles*. 1992;34(11). doi:10,3987/COM-92-6142.

31. Watanabe T, Mazumder TK, Yamamoto A, et al. A simple and rapid method for analyzing the Monascus pigment-mediated degradaton of mutagenic 3-hydroxyamino-1-methyl-5H-pyrido[4,3-b]indole by in-capillary micellar electrokinetic chromatography. *Mutat Res*. 1999;444:75-83.

32. Li YG, Zhang F, Wang ZT, Hu ZB. Identification and chemical profiling of monacolins in red yeast rice using high-performance liquid chromatography with photodiode array detector and mass spectrometry. *J Pharm Biomed Anal*. 2004;35(5):1101-1112. doi:10.1016/j.jpba.2004.04.004.

33. Juzlova P, Martinkova L, Kfen V. Secondary metabolites of the fungus Monascus : a review. *J Ind Microbiol Biotechnol*. 1996;16(April 1995):163-170.

34. Blanc PJ. Characterization of monascidin A from Monascus as citrinin. *Int J Food Microbiol*. 1995;27:201-213.

35. Blanc PJ, Loret MO, Goma G, Rangueil CS De. Production of citrinin by various speecies of monascus. *Biotechnol Lett*. 1995;17(3):291-294.

36. Hajjaj H, Blanc P, Groussac E, Goma G, Uribelarrea J, Loubiere P. Improvement of red pigment/citrinin production ratio as a function of environmental conditions by monascus ruber. *Biotechnol Bioeng*.

1999;64(4):497-501.

37. Zhang L, Li Z, Dai B, Zhang W, Yuan Y. Effect of submerged and solid-state fermentation on pigment and citrinin production by Monascus purpureus. *Acta Biol Hung*. 2013;64(3):385-394. doi:10.1556/ABiol.64.2013.3.11.

38. Wang Y-Z, Ju X-L, Zhou Y-G. The variability of citrinin production in Monascus type cultures. *Food Microbiol*. 2005;22(1):145-148. doi:10.1016/j.fm.2004.01.006.

39. Kang B, Zhang X, Wu Z, Wang Z, Park S. Production of citrinin-free Monascus pigments by submerged culture at low pH. *Enzyme Microb Technol*. 2014;55:50-57. doi:10.1016/j.enzmictec.2013.12.007.

40. Chen D, Xue Y, Chen M, Li Z, Wang C. Optimization of submerged fermentation medium for citrinin-free monascin production by *Monascus*. *Prep Biochem Biotechnol*. 2016;6068(April):10826068.2015.1135461. doi:10.1080/10826068.2015.1135461.

41. Ogihara J, Fujimotoz Y. Biosynthesis of PP-V, a Monascorubramine Homologue by Penicillium sp. AZ. *J Biosci Bioeng*. 2000;90(6):2-4.

42. Woo PCY, Lam C-W, Tam EWT, et al. The biosynthetic pathway for a thousand-year-old natural food colorant and citrinin in Penicillium marneffei. *Sci Rep*. 2014;4:6728. doi:10.1038/srep06728.

43. Hajjaj H, Klaébé A, Loret MO, Goma G, Blanc PJ, François J. Biosynthetic Pathway of Citrinin in the Filamentous Fungus Monascus ruber as Revealed by 13C Nuclear Magnetic Resonance. *Appl Environ Microbiol*. 1999;65(1):311-314.

44. Balakrishnan B, Karki S, Chiu SH, et al. Genetic localization and in vivo characterization of a Monascus azaphilone pigment biosynthetic gene cluster. *Appl Microbiol Biotechnol*. 2013;97(14):6337-6345. doi:10.1007/s00253-013-4745-9.

45. Balakrishnan B, Chen CC, Pan TM, Kwon HJ. Mpp7 controls regioselective Knoevenagel condensation during the biosynthesis of Monascus azaphilone pigments. *Tetrahedron Lett*. 2014;55(9):1640-1643. doi:10.1016/j.tetlet.2014.01.090.

46. Bijinu B, Suh J-W, Park S-H, Kwon H-J. Delineating Monascus azaphilone pigment biosynthesis: oxidoreductive modifications determine the ring cyclization pattern in azaphilone biosynthesis. *RSC Adv*. 2014;4(103):59405-59408. doi:10.1039/C4RA11713A.

47. Lin TF, Yakushijin K, Buchi GH, Demain AL, Bt GH, Demain AL. Formation of water-soluble Monascus red pigments by biological and semi-synthetic processes. *J Ind Microbiol*. 1992;9(3-4):173-179. doi:10.1007/BF01569621.

48. Ogihara J, Oishi K. Effect of ammonium nitrate on the production of PP-V and monascorubrin homologues by Penicillium sp. AZ. *J Biosci Bioeng*. 2002;93(1):54-59. doi:10.1016/S1389-1723(02)80054-8.

49. Blanc PJ, Loret MO, Santerre a L, et al. Pigments of Monascus. *J Food Sci*. 1994;59(4):862-865. doi:10.1111/j.1365-2621.1994.tb08145.x.

50. Ogihara J. Effect of Ammonium Nitrate on the Production of PP-V and. 2001;93(1):54-59.

51. Sato. Identification of major pigments containing D-amino acid units in commercial Monascus pigments. *ChemPharmBull*. 1997;45(1):227-229. doi:10.1248/cpb.37.3229.

52. Jung H, Kim C, Kim K, Shin CS. Color characteristics of monascus pigments derived by fermentation with various amino acids. *J Agric Food Chem*. 2003;51(5):1302-1306. doi:10.1021/jf0209387.

53. Jung H, Kim C, Shin CS. Enhanced photostability of monascus pigments derived with various amino acids via fermentation. *J Agric Food Chem*. 2005;53(18):7108-7114. doi:10.1021/jf0510283.

54. Jo D, Deokyeong C, Nam K, Shin CS. Biological evaluation of novel derivatives of the orange pigments from Monascus sp. as inhibitors of melanogenesis. *Biotechnol Lett*. 2014;36(8):1605-1613. doi:10.1007/s10529-014-1518-1.

55. Xiong X, Zhang X, Wu Z, Wang Z. Accumulation of yellow Monascus pigments by extractive fermentation in nonionic surfactant micelle aqueous solution. *Appl Microbiol Biotechnol*. 2014;99(3):1173-1180. doi:10.1007/s00253-014-6227-0.

56. Workman M, Andersen MR, Thykaer J. Integrated Approaches for Assessment of Cellular Performance in Industrially Relevant Filamentous Fungi. *Ind Biotechnol*. 2013;9(6):337-344. doi:10.1089/ind.2013.0025.

57. Mapari S. Production of monascus-like azaphilone pigment.

58. Frisvad JC, Filtenborg O, Samson R a, Stolk a C. Chemotaxonomy of the genus Talaromyces. *Antonie Van Leeuwenhoek*. 1990;57(3):179-189.

59. Frisvad JC, Yilmaz N, Thrane U, Rasmussen KB, Houbraken J, Samson R a. Talaromyces atroroseus, a new species efficiently producing industrially relevant red pigments. *PLoS One*. 2013;8(12):1-15. doi:10.1371/journal.pone.0084102.

60. Arai T, Umemura S, Ota T, Ogihara J, Kato J, Kasumi T. Effects of Inorganic Nitrogen Sources on the Production of PP-V [(10Z)-12-carboxyl-monascorubramine] and the Expression of the Nitrate Assimilation Gene Cluster by Penicillium sp. AZ. *Biosci Biotechnol Biochem*. 2012;76(1):120-124. doi:10.1271/bbb.110589.

61. Arai T, Koganei K, Umemura S, et al. Importance of the ammonia assimilation by Penicillium purpurogenum in amino derivative Monascus pigment, PP-V, production. *AMB Express*. 2013;3(1):19. doi:10.1186/2191-0855-3-19.

62. Arai T, Kojima R, Motegi Y, Kato J, Kasumi T, Ogihara J. PP-O and PP-V, Monascus pigment homologues, production, and phylogenetic analysis in Penicillium purpurogenum. *Fungal Biol*. 2015;119(12):1226-1236. doi:10.1016/j.funbio.2015.08.020.

63. Ogihara J, Kato J, Oishi K, Fujimoto Y, Ogihara. PP-R, 7-(2-hydroxyethyl)-monascorubramine, a red pigment produced in the mycelia of Penicillium sp. AZ. *J Biosci Bioeng*. 2000;91(1):44-47. doi:10.1016/S1389-1723(01)80109-2.

64. Ogihara J, Kato J, Oishi K, Fujimoto Y, Eguchi T. Production and structural analysis of PP-V, a homologue of monascorubramine, produced by a new isolate of Penicillium sp. *J Biosci Bioeng*. 2000;90(5):549-554.

65. Mapari SA, Meyer AS, Thrane U, Frisvad JC. Identification of potentially safe promising fungal cell factories for the production of polyketide natural food colorants using chemotaxonomic rationale. *Microb Cell Fact*. 2009;8:24. doi:10.1186/1475-2859-8-24.

66. Mapari S a S, Meyer AS, Thrane U. Colorimetric characterization for comparative analysis of fungal pigments and natural food colorants. *J Agric Food Chem*. 2006;54(19):7027-7035. doi:10.1021/jf062094n.

67. King T. Studies in Mycological Chemistry. Part XXX and Last. Isolation and Structure of Purpuride, a Metabolite of Penicillium purpurogenum Stoll. 1973:78-80.

68. Ventura SPM, Santos-Ebinuma VC, Pereira JFB, Teixeira MFS, Pessoa A, Coutinho JAP. Isolation of

natural red colorants from fermented broth using ionic liquid-based aqueous two-phase systems. *J Ind Microbiol Biotechnol*. 2013;40(5):507-516. doi:10.1007/s10295-013-1237-y.

69. Santos-Ebinuma VC, Teixeira MFS, Pessoa A. Submerged culture conditions for the production of alternative natural colorants by a new isolated Penicillium purpurogenum DPUA 1275. *J Microbiol Biotechnol*. 2013;23(6):802-810.

70. Santos-Ebinuma VC, Roberto IC, Simas Teixeira MF, Pessoa A. Improving of red colorants production by a new Penicillium purpurogenum strain in submerged culture and the effect of different parameters in their stability. *Biotechnol Prog*. 2013;29(3):778-785. doi:10.1002/btpr.1720.

71. Santos-Ebinuma VC, Roberto IC, Francisca M, et al. Improvement of submerged culture conditions to produce colorants by Penicillium purpurogenum. *Brazilian J Microbiol*. 2014;742:731-742.

72. Méndez A, Pérez C, Montañéz JC, Martínez G, Aguilar CN. Red pigment production by Penicillium purpurogenum GH2 is influenced by pH and temperature. *J Zhejiang Univ Sci B*. 2011;12(12):961-968. doi:10.1631/jzus.B1100039.

73. Gunasekaran S, Poorniammal R. Optimization of fermentation conditions for red pigment production from Penicillium sp . under submerged cultivation. *African J Biotechnol*. 2008;7(12):1894-1898.

74. Mapari S a S, Hansen ME, Meyer AS, Thrane U. Computerized screening for novel producers of Monascus-like food pigments in Penicillium species. *J Agric Food Chem*. 2008;56(21):9981-9989. doi:10.1021/jf801817q.

75. Tam EWT, Tsang CC, Lau SKP, Woo PCY. Polyketides, toxins and pigments in Penicillium marneffei. *Toxins (Basel)*. 2015;7(11):4421-4436. doi:10.3390/toxins7114421.

76. Velmurugan P, Lee YH, Venil CK, Lakshmanaperumalsamy P, Chae J-C, Oh B-T. Effect of light on growth, intracellular and extracellular pigment production by five pigment-producing filamentous fungi in synthetic medium. *J Biosci Bioeng*. 2010;109(4):346-350. doi:10.1016/j.jbiosc.2009.10.003.

77. Broder CU, Koehler PE. PIGMENTS PRODUCED BY Monascus purpureus WITH REGARD TO QUALITY AND QUANTITY. *Inst Food Technol*. 1980;45(1973):567-569. doi:10.1111/j.1365-2621.1980.tb04102.x.

78. Babitha S, Soccol CR, Pandey A. Solid-state fermentation for the production of Monascus pigments from jackfruit seed. *Bioresour Technol*. 2007;98(8):1554-1560. doi:10.1016/j.biortech.2006.06.005.

79. Hu Z, Zhang X, Wu Z, Qi H, Wang Z. Perstraction of intracellular pigments by submerged cultivation of Monascus in nonionic surfactant micelle aqueous solution. *Appl Microbiol Biotechnol*. 2012;94(1):81-89. doi:10.1007/s00253-011-3851-9.

80. Kongruang S. Growth kinetics of biopigment production by Thai isolated Monascus purpureus in a stirred tank bioreactor. *J Ind Microbiol Biotechnol*. 2011;38(1):93-99. doi:10.1007/s10295-010-0834-2.

81. Mohamed MS, Mohamad R, Manan MA, Ariff AB. Enhancement of Red Pigment Production by Monascus purpureus FTC 5391 through Retrofitting of Helical Ribbon Impeller in Stirred-Tank Fermenter. *Food Bioprocess Technol*. 2012;5(1):80-91. doi:10.1007/s11947-009-0271-2.

82. Mapari S a S, Meyer AS, Thrane U. Photostability of natural orange-red and yellow fungal pigments in liquid food model systems. *J Agric Food Chem*. 2009;57(14):6253-6261. doi:10.1021/jf900113q.

83. Zhou B, Wang J, Pu Y, Zhu M, Liu S, Liang S. Optimization of culture medium for yellow pigments production with Monascus anka mutant using response surface methodology. *Eur Food Res Technol*. 2009;228(6):895-901. doi:10.1007/s00217-008-1002-z.

84. Kim C, Jung H, Kim YO, Shin CS. Antimicrobial activities of amino acid derivatives of monascus pigments. *FEMS Microbiol Lett*. 2006;264(1):117-124. doi:10.1111/j.1574-6968.2006.00451.x.

85. Sun X, Yang X, Wang E. Chromatographic and electrophoretic procedures for analyzing plant pigments of pharmacologically interests. *Anal Chim Acta*. 2005;547(2):153-157. doi:10.1016/j.aca.2005.05.051.

86. Watanabe T, YAMAMOTO A. Separation and determination of Monascus Yellowfor food by Micellar Electrokinetic Chromatography Pigments. *Anal Sci*. 1997;13(August).

87. Shi K, Song D, Chen G, Pistolozzi M, Wu Z, Quan L. Controlling composition and color characteristics of Monascus pigments by pH and nitrogen sources in submerged fermentation. *J Biosci Bioeng*. 2015;xx(xx). doi:10.1016/j.jbiosc.2015.01.001.

88. Hajjaj H, Klaébé A, Goma G, Blanc PJ, Barbier E, François J. Medium-chain fatty acids affect citrinin production in the filamentous fungus Monascus ruber. *Appl Environ Microbiol*. 2000;66(3):1120-1125. doi:10.1128/AEM.66.3.1120-1125.2000.

89. Zhong S, Zhang X, Wang Z. Preparation and characterization of yellow Monascus pigments. *Sep Purif Technol*. 2015;150:139-144. doi:10.1016/j.seppur.2015.06.040.

90. Panfili G, Fratianni A, Irano M. Improved normal-phase high-performance liquid chromatography procedure for the determination of carotenoids in cereals. *J Agric Food Chem*. 2004;52(21):6373-6377. doi:10.1021/jf0402025.

91. Hsu YW, Hsu LC, Liang YH, Kuo YH, Pan TM. New bioactive orange pigments with yellow fluorescence from Monascus-fermented dioscorea. *J Agric Food Chem*. 2011;59(9):4512-4518. doi:10.1021/jf1045987.

92. Kildgaard S, Mansson M, Dosen I, et al. Accurate dereplication of bioactive secondary metabolites from marine-derived fungi by UHPLC-DAD-QTOFMS and a MS/HRMS library. *Mar Drugs*. 2014;12(6):3681-3705. doi:10.3390/md12063681.

93. Domenici V, Ancora D, Cifelli M, Serani A, Veracini CA, Zandomeneghi M. Extraction of pigment information from near-UV vis absorption spectra of extra virgin olive oils. *J Agric Food Chem*. 2014;62(38):9317-9325. doi:10.1021/jf503818k.

94. Torres PB, Chow F, Furlan CM, Mandelli F, Mercadante A, dos Santos DYAC. Standardization of a protocol to extract and analyze chlorophyll a and carotenoids in Gracilaria tenuistipitata var. liui. zhang and xia (rhodophyta). *Brazilian J Oceanogr*. 2014;62(1):57-63. doi:10.1590/S1679-87592014068106201.

95. Said FM, Brooks J, Chisti Y. Optimal C:N ratio for the production of red pigments by Monascus ruber. *World J Microbiol Biotechnol*. 2014;30(9):2471-2479. doi:10.1007/s11274-014-1672-6.

96. Fu W, Magnúsdóttir M, Brynjólfson S, Palsson B, Paglia G. UPLC-UV-MSE analysis for quantification and identification of major carotenoid and chlorophyll species in algae. *Anal Bioanal Chem*. 2012;404(10):3145-3154. doi:10.1007/s00216-012-6434-4.

97. Escolar D, Haro MR, Saucedo A, Ayuso J, Jiménez A, Alvarez JA. Measurement of the concentrations of solution through chromatic systems. *Appl Opt*. 1995;34(19)(19):3731-3735.

98. Mørkeberg R, Carlsen M, Nielsen J. Induction and repression of alpha-amylase production in batch and continuous cultures of Aspergillus oryzae. *Microbiology*. 1995;141 (Pt 1(1 995):2449-2454. doi:10.1002/(SICI)1097-0290(19960205)49:3<266::AID-BIT4>3.0.CO;2-I.

99. Yang J, Chen Q, Wang W, Hu J, Hu C. Effect of oxygen supply on Monascus pigments and citrinin

production in submerged fermentation. *J Biosci Bioeng*. 2015;119(5):564-569. doi:10.1016/j.jbiosc.2014.10.014.

100. Pedersen H, Beyer M, Nielsen J. Glucoamylase production in batch, chemostat and fed-batch cultivations by an industrial strain of Aspergillus niger. *Appl Microbiol Biotechnol*. 2000;53(3):272-277. doi:10.1007/s002530050020.

101. Babitha S, Soccol CR, Pandey A. Effect of stress on growth, pigment production and morphology of Monascus sp. in solid cultures. *J Basic Microbiol*. 2007;47(2):118-126. doi:10.1002/jobm.200610261.

102. Rajendhran J, Krishnakumar V, Gunasekaran P. Optimization of a fermentation medium for the production of Penicillin G acylase from Bacillus sp. *Lett Appl Microbiol*. 2002;35(6):523-527. doi:10.1046/j.1472-765X.2002.01234.x.

103. Yoshimura M. Production of Monascus-Pigment in a Submerged Cultures. 1975;39(9):1789-1795.

104. Prajapati VS, Soni N, Trivedi UB, Patel KC. An enhancement of red pigment production by submerged culture of Monascus purpureus MTCC 410 employing statistical methodology. *Biocatal Agric Biotechnol*. 2014;3(2):140-145. doi:10.1016/j.bcab.2013.08.008.

105. Mapari S. Prodcution of Monascus-like azaphilone pigments. 2010.

106. Krairak S, Yamamura K, Irie R, et al. Maximizing yellow pigment production in fed-batch culture of Monascus sp. *J Biosci Bioeng*. 2000;90(4):363-367. doi:10.1016/S1389-1723(01)80002-5.

107. Chen M, Johns MR. Effect of pH and nitrogen source on pigment production by Monascus purpureus. *Appl Microbiol Biotechnol*. 1993:132-138.

108. Lin TF, Demain AL. Effect of nutrition of Monascus sp. on formation of red pigments. 1991:70-75.

109. Koehler BPE. PIGMENTS PRODUCED BY Monascus purpureus WITH REGARD TO QUALITY AND QUANTITY. 1980;45(1973):567-569.

110. Steiner J, Socha C, Eyzaguirre J. Culture conditions for enhanced cellulase production by a native strain of Penicillium purpurogenum. *World J Microbiol …*. 1994;10(562):280-284.

111. Yang Y, Liu B, Du X, et al. Complete genome sequence and transcriptomics analyses reveal pigment biosynthesis and regulatory mechanisms in an industrial strain, Monascus purpureus YY-1. *Sci Rep*. 2015;5:8331. doi:10.1038/srep08331.

112. Juzlova P. Ethanol as substrate for pigment production by the fungus Monascus purpureus. 1994:7-12.

113. Lin TF, Demain AL. Leucine interference in the produc- tion of water-soluble red Monascus pigments. *arch microbiol*. 1994;162:114-119.

114. Carels M, Shepherd D. The effect of different nitrogen sources on pigment production and sporulation of Monascus species in submerged, shaken culture. *Can J Microbiol*. 1977;23(10):1360-1372. http://www.ncbi.nlm.nih.gov/pubmed/21736.

115. Lin TF, Demain AL. Negative effect of ammonium nitrate as nitrogen source on the production of water-soluble red pigments by Monascus sp. 1994:701-705.

116. Carels M, Shepherd D. The effect of pH and amino acids on conidiation and pigment production of Monascus major ATCC 16362 and Monascus rubiginosus ATCC 16367 in submerged shaken culture. *Can J Microbiol*. 1978;24(11):1346-1357. http://www.ncbi.nlm.nih.gov/pubmed/33756.

117. Lin TF, Demain AL. Resting cell studies on formation of water-soluble red pigments by Monascus sp .

1993;12:361-367.

118. Orozco SFB, Kilikian BV. Effect of pH on citrinin and red pigments production by Monascus purpureus CCT3802. *World J Microbiol Biotechnol*. 2008;24(2):263-268. doi:10.1007/s11274-007-9465-9.

119. Joshi VK, Attri D, Baja A, Bhushan S. Microbial Pigments. 2003;2(July):362-369.

120. Ahn J, Jung J, Hyung W, Haam S, Shin C. Enhancement of monascus pigment production by the culture of Monascus sp. J101 at low temperature. *Biotechnol Prog*. 2006;22(1):338-340. doi:10.1021/bp050275o.

121. Han, Ohantaek, Mudgett RE. Effects of Oxygen and Carbon Dioxide Partial Pressures on. 1992:5-10.

122. Pereira DG, Tonso A, Kilikian B V. Effect of dissolved oxygen concentration on red pigment and citrinin production by Monascus purpureus ATCC 36928. *Brazilian J Chem Eng*. 2008;25(2):247-253. doi:10.1590/S0104-66322008000200004.

123. Bühler RMM, Müller BL, Moritz DE, Vendruscolo F, de Oliveira D, Ninow JL. Influence of Light Intensity on Growth and Pigment Production by Monascus ruber in Submerged Fermentation. *Appl Biochem Biotechnol*. 2015:1277-1289. doi:10.1007/s12010-015-1645-8.

124. Babitha S, Carvahlo JC, Soccol CR, Pandey A. Effect of light on growth, pigment production and culture morphology of Monascus purpureus in solid-state fermentation. *World J Microbiol Biotechnol*. 2008;24(11):2671-2675. doi:10.1007/s11274-008-9794-3.

125. Miyake T, Mori A, Kii T, et al. Light effects on cell development and secondary metabolism in Monascus. *J Ind Microbiol Biotechnol*. 2005;32(3):103-108. doi:10.1007/s10295-005-0209-2.

126. Johns MR, M SD. Production of pigments by Monascus purpureus in solid culture. *J Ind Microbiol*. 1991;2007:23-28.

127. Carvalho JC, Oishi BO, Woiciechowski AL, Pandey A, Babitha S, Soccol CR. Effect of substrates on the production of Monascus biopigments by solid-state fermentation and pigment extraction using different solvents. *Indian J Biotechnol*. 2007;6(2):194-199.

128. Liu DC, Wu SW, Tan FJ. Effects of addition of anka rice on the qualities of low-nitrite Chinese sausages. *Food Chem*. 2010;118(2):245-250. doi:10.1016/j.foodchem.2009.04.114.

129. Nimnoi P, Lumyong S. Improving Solid-State Fermentation of Monascus purpureus on Agricultural Products for Pigment Production. *Food Bioprocess Technol*. 2011;4(8):1384-1390. doi:10.1007/s11947-009-0233-8.

130. Lotong N, Suwanarit P. Fermentation of ang-kak in plastic bags and regulation of pigmentation by initial moisture content. *J Appl Microbiol*. 1990.

131. Vendruscolo F, Bühler RMM, de Carvalho JC, et al. Monascus: a Reality on the Production and Application of Microbial Pigments. *Appl Biochem Biotechnol*. 2015:211-223. doi:10.1007/s12010-015-1880-z.

132. Christensen LH, Henriksen CM. Continuous cultivation of Penicillium chrysogenum. Growth on glucose and penicillin production. *J Biotechnol*. 1995;42:95-107.

133. Shao Y, Lei M, Mao Z, Zhou Y, Chen F. Insights into Monascus biology at the genetic level. *Appl Microbiol Biotechnol*. 2014;98(9):3911-3922. doi:10.1007/s00253-014-5608-8.

134. Chen Y, Tseng C, Liaw L, et al. Cloning and Characterization of Monacolin K Biosynthetic Gene Cluster

from Monascus pilosus Cloning and Characterization of Monacolin K Biosynthetic Gene Cluster from Monascus pilosus. *Society*. 2008:5639-5646. doi:10.1021/jf800595k.

135. Shimizu T, Kinoshita H, Ishihara S, Sakai K, Nagai S. Polyketide Synthase Gene Responsible for Citrinin Biosynthesis in Monascus purpureus Polyketide Synthase Gene Responsible for Citrinin Biosynthesis in Monascus purpureus. 2005;71(7):1-6. doi:10.1128/AEM.71.7.3453.

136. He Y, Liu Q, Shao Y, Chen F. Ku70 and ku80 null mutants improve the gene targeting frequency in Monascus ruber M7. *Appl Microbiol Biotechnol*. 2013;97(11):4965-4976. doi:10.1007/s00253-013-4851-8.

137. Davison J, al Fahad A, Cai M, et al. Genetic, molecular, and biochemical basis of fungal tropolone biosynthesis. *Proc Natl Acad Sci*. 2012;109(20):7642-7647. doi:10.1073/pnas.1201469109.

138. Joardar V, Abrams NF, Hostetler J, et al. Sequencing of mitochondrial genomes of nine Aspergillus and Penicillium species identifies mobile introns and accessory genes as main sources of genome size variability. *BMC Genomics*. 2012;13(1):698. doi:10.1186/1471-2164-13-698.

139. Woo PCY, Tam EWT, Chong KTK, et al. High diversity of polyketide synthase genes and the melanin biosynthesis gene cluster in Penicillium marneffei. *FEBS J*. 2010;277:3750-3758. doi:10.1111/j.1742-4658.2010.07776.x.

140. Nødvig CS, Nielsen JB, Kogle ME, Mortensen UH. A CRISPR-Cas9 system for genetic engineering of filamentous fungi. *PLoS One*. 2015;10(7):1-18. doi:10.1371/journal.pone.0133085.

141. Meyer V, Arentshorst M, El-Ghezal A, et al. Highly efficient gene targeting in the Aspergillus niger kusA mutant. *J Biotechnol*. 2007;128(4):770-775. doi:10.1016/j.jbiotec.2006.12.021.

142. Su NW, Lin YL, Lee MH, Ho CY. Ankaflavin from Monascus-fermented red rice exhibits selective cytotoxic effect and induces cell death on Hep G2 cells. *J Agric Food Chem*. 2005;53(6):1949-1954. doi:10.1021/jf048310e.

143. Wong H, Koehler P. Production of Red Water-soluble Monascus Pigments. *J Food Sci*. 1983;48.

144. Gheith O, Sheashaa H, Abdelsalam M. Efficacy and safety of Monascus purpureus Went rice in children and young adults with secondary hyperlipidemia: A preliminary report. *Eur J Intern Med*. 2008. doi:http://dx.doi.org/10.1016/j.ejim.2008.08.012.

145. Bianchi A. Extracts of Monascusus purpureus beyond statins-profile of efficacy and safety of the use of extracts of Monascus purpureus. *Chin J Integr Med*. 2005;11(4):309-313.

146. Hsu WH, Pan TM. Treatment of metabolic syndrome with ankaflavin, a secondary metabolite isolated from the edible fungus Monascus spp. *Appl Microbiol Biotechnol*. 2014;98(11):4853-4863. doi:10.1007/s00253-014-5716-5.

147. Mohan Kumari HP, Akhilender Naidu K, Vishwanatha S, Narasimhamurthy K, Vijayalakshmi G. Safety evaluation of Monascus purpureus red mould rice in albino rats. *Food Chem Toxicol*. 2009;47(8):1739-1746. doi:10.1016/j.fct.2009.04.038.

148. Yu CC, Wang JJ, Lee CL, Lee SH, Pan TM. Safety and mutagenicity evaluation of nanoparticulate red mold rice. *J Agric Food Chem*. 2008;56(22):11038-11048. doi:10.1021/jf801335u.

149. Osmanova N, Schultze W, Ayoub N. Azaphilones: a class of fungal metabolites with diverse biological activities. *Phytochem Rev*. 2010:1-28. doi:10.1007/s11101-010-9171-3.

150. Akihisa T, Tokuda H, Yasukawa K, et al. Azaphilones, furanoisophthalides, and amino acids from the

extracts of Monascus pilosus-fermented rice (red-mold rice) and their chemopreventive effects. *J Agric Food Chem*. 2005;53(3):562-565. doi:10.1021/jf040199p.

151. Jang H, Choe D, Shin CS. Novel derivatives of monascus pigment having a high CETP inhibitory activity. *Nat Prod Res*. 2014;28(18):1427-1431. doi:10.1080/14786419.2014.905561.

152. Wong H-C, Bau Y-S. Pigmentation and Antibacterial Activity of Fast Neutron- and X-Ray-induced Strains of Monascus purpureus Went. *Plant Physiol*. 1977;60:578-581.

153. Martinkova L, Patakova-Juzlova P, Krent V, et al. Biological activities of oligoketide pigments of Monascus purpureus. *Food Addit Contam*. 1999;16(1):15-24. doi:10.1080/026520399284280.

154. Lee CL, Kung YH, Wu CL, Hsu YW, Pan TM. Monascin and ankaflavin act as novel hypolipidemic and high-density lipoprotein cholesterol-raising agents in red mold dioscorea. *J Agric Food Chem*. 2010;58(16):9013-9019. doi:10.1021/jf101982v.

155. Kim JH, KIM YO, JEUN J, CHOI D-Y, SHIN CS. Trp and Leu-OEt Derivatives of the Monascus Pigment Exert High Anti-Obesity Effects on Mice. *Biosci Biotechnol Biochem*. 2010;74(2):304-308. doi:10.1271/bbb.90620.

Appendix 4

New azaphilones from *Aspergillus neoglaber*

Thomas Isbrandt and Thomas Ostenfeld Larsen

Department of Biotechnology and Biomedicine, Technical University of Denmark, Kongens Lyngby, Denmark

Abstract

Three new azaphilones, sassafrin E, sassafrin F, and sassafrinamine A, were isolated from the filamentous fungus *Aspergillus neoglaber*. The structure of the compounds were determined by nuclear magnetic resonance spectroscopy, and were found to be novel analogues of two already known classes, sassfrins and berkchaetoazaphilones. Sassafrin E and F were both oxygen containing, while sassafrinamine contained a nitrogen atom, originating from an ethanolamine moiety, as well as extensive conjugation resulting in an intense purple colour of the pure compound. The structrure of sassafrin E was further confirmed using deuterium exchange experiments coupled with high resolution tandem mass spectrometry.

Introduction

Discovery of novel microbial secondary metabolites with bioactive properties has for decades been an important research field since a large proportion of new pharmaceutical are derived from natural products.[1] One strategy for finding new compounds is by looking into the secondary metabolites of newly described species. The secondary metabolites of the filamentous fungus *Aspergillus neoglaber* (= *Neosartorya glabra*) has not been well described, and only little research has been done to gain a better understanding of which compounds are produced by this species.[2–4]. Some of the known secondary metabolites include the bioactive glabramycins[3,5] and satoryglabrins[4], and various diketopiperazines and tetracyclopeptides[2]. Investigation of the secondary metabolite profile by HPLC-MS tentatively suggested the presence of additional already known bioactive secondary metabolite, among these the acetylcholineesterase inhibitor, arisugacin C, previously identified in *Penicillium echinulatum*.[6] The prospect of identifying novel analouges belonging to this compounds class incited us to further study the secondary metabolite capabilities of *A. neoglaber*.

Results and Discussion

Ultra-high performance liquid chromatography coupled to diode array detection and high resolution tandem mass spectrometry (UHPLC-DAD-HRMS/MS) analysis of the ethyl acetate extract (Figure 1) from the filamentous fungus *Aspergillus neoglaber* IBT3020 led to the identification of a compound (**1**) with the same accurate mass and molecular formula (m/z 453.2277, $[M+H]^+$, $C_{27}H_{32}O_6$) as the acetylcholineesterase inhibitor

arisugacin C.[6,7] Further investigation by comparison of retention time, UV-VIS spectrum, and fragmentation pattern, with an extract from *Penicillium echinulatum*, a known producer of arisugacin C, indicated that the compound in *A. neoglaber* was not arisugacin C, but rather an analogue or an entirely different secondary metabolite (see supplementary material S1). In addition to **1**, a second compound (**2**) with a mass and molecular formula corresponding to addition of two protons was tentatively identified as a likely analogue (m/z 455.2428, $[M+H]^+$, $C_{27}H_{34}O_6$). The two compounds were purified along with a third compound (**3**) absorbing at 550 nm with a m/z of 494.2541 (molecular formula $C_{29}H_{36}NO_6$).

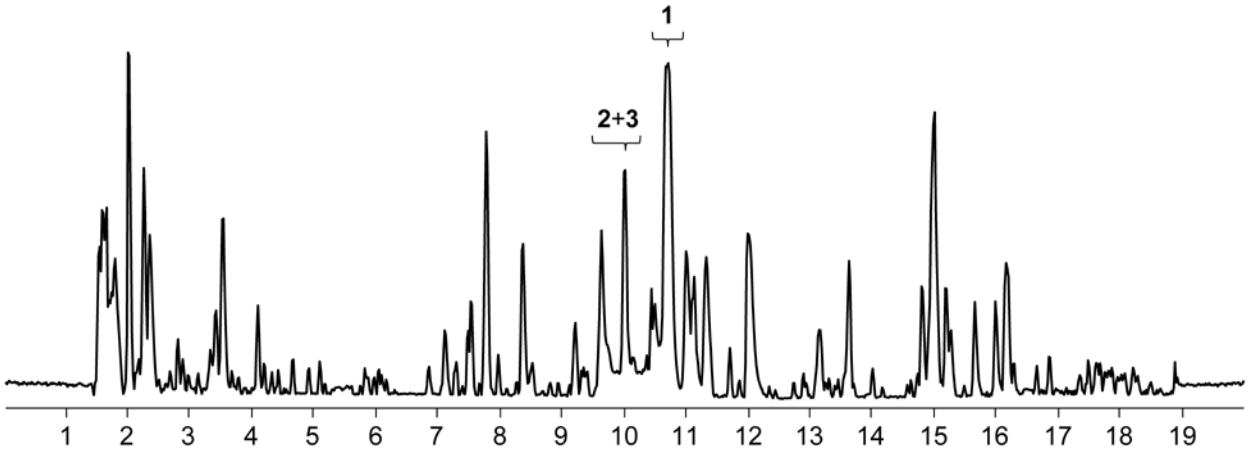


Figure 1. Base peak chromatogram of the ethyl acetate extract from A. neoglaber grown on YES medium.

In order to purify the three compounds, *A. neoglaber* IBT 3020 was cultivated on 6x500mL semi liquid YES media. The biomass was extracted with ethyl acetate, and purification was done using normal phase flash chromatography followed by semi-preparative RP-HPLC. One- and two-dimensional NMR experiments were used in order to elucidate the structure of the compounds.

In compound **1**, a total of 27 protons could be identified from the ^{1}H-spectrum, matching the expected formula. In combination with edHSQC, 12 CH-groups, two CH_2-groups, and five CH_3-groups could be identified. Eight of the CH groups had carbon shifts matching alkenes, and one was identified to be attached to a hydroxyl group. Both CH_2-groups appeared as diastereotropic. ^3J H-H couplings obtained from DQF-COSY, identified four spin systems consisting of H-1 to H-3, H-13 to H-16 and H-15-CH_3, H-18 to H-21, and H-10 and H-23, as well as five singlets. Correlations in the DQF-COSY, were all confirmed in H2BC. HMBC correlations linked H-15, H-17-CH_3 and H-18 to C-17, and H-10 and H-20 to C-22. Additionally, H-1, H-2 and H-3, along with H-5 was linked to C-4. Ambiguity in the HMBC around C-6 and C-11 meant that additional, more specific experiments were needed, and 1,n- and 1,1-ADEQUATE were able to connect H-5 and H-7 to C-6, H-10 to C-11, as well as H-9-CH_3 to C-24 (Figure 2). NOESY correlations around the lactone could assign relative stereochemistry to the methyl group on C-9 and the two protons H-10 and H-23. In summary,

compound **1** turned out to be a novel azaphilone, with high structural similarities with groups of compounds such as sassafrins[8] and berkchaetoazaphilones[9], and have been named sassafrin E.

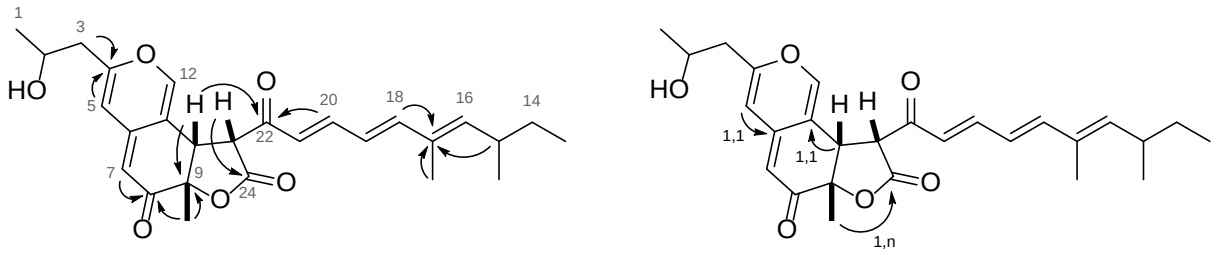


*Figure 2. Structure and key HMBC and ADEQUATE correlations for sassafrin E (**1**).*

The NMR data for compounds **2** and **3**, was highly similar to that of **1**, with only few variations. Compound **2** was determined to only differ from **1**, by having the ketone at C-22 being reduced to a hydroxyl group and has been named sassafrin F.

*Table 1. ^{1}H and ^{13}C NMR shifts for compounds **1, 2,** and **3**.*

	Sassafrin E (**1**)		Sassafrin F (**2**)		Sassafrinamine A (**3**)	
#	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C
1	1.22	23.4	1.23	22.1	1.33	23.6
2	4.06	66.2	40.01	65.3	4.15	67.9
3	2.52/2.59	43.7	2.55	42.3	2.95/3.01	41.5
4	-	162.9	-	161.3	-	154.7
5	6.34	109.9	6.34	108.5	7.08	122.4
6	-	148.2	-	147.3	-	?
7	5.41	105.7	5.38	104	6.86	99.1
8	-	194.1	-	193.3	-	196
9	-	84.3	-	82.7	-	86.5
9-CH3	1.55	23.5	1.55	21.5	1.69	30.1
10	4.03	44.6	3.63	42.1	-	?
11	-	116.4	-	114.8	-	119.9
12	7.53	149.8	7.42	150	8.35	144.4
13	0.86	12.3	0.87	10.9	0.87	12
14	1.30/1.42	31.2	1.29/1.42	30.1	1.31/1.43	31
15	2.48	36.1	2.45	34.3	2.48	35.6
15-CH3	0.99	20.6	0.99	19.6	0.99	20.5
16	5.60	147.4	5.32	139.7	5.51	144.9
17	-	134.4	-	138.1	-	134.2
17-CH3	1.82	12.6	1.78	11.4	1.85	12.4
18	6.77	150.4	6.3	138.1	6.68	146.9
19	6.40	125.7	6.15	125	6.45	126.4
20	7.41	148.4	6.4	131.1	7.36	142.4
21	6.46	128.2	5.65	132.4	7.55	127.9
22	-	192.4	4.83	68.1	-	185.8
23	4.41	55.7	3.01	48.7	-	?
24	-	171.3	-	174.4	-	173.6
1'					4.32/4.51	57.7
2'					3.89	61.3

Compound **3** (m/z 494.2541, $[M+H]^+$, $C_{29}H_{36}NO_6$) included three additional hydrogen atoms, two more carbon atoms, as well as a nitrogen, relative to **1**. From HSQC, the additional carbon atoms were determined to be two CH_2-groups (C-1' and C-2'). HMBC correlations from H-12 to C-1' determined the two-carbon moiety to be N-linked to the isoquinoline ring. Furthermore, no signals were observed for protons H-10 and H-23, and it is therefore assumed that these are connected via a double bond. Similarly, no correlations to C-6 were observed. The UV-VIS spectrum for **3** was quite unique, with slight absorption all the way from 270 nm to 580nm, with maximum at 545 nm (SUPP), and the extensive conjugation agrees well with the violet/purple colour of the pure compound. Compound **3** has been named sassarinamine A[1], based on the incorporation of nitrogen. Structures of compounds **1**, **2**, and **3** are shown in Figure 3 and chemical shifts are listed in Table 1. Recorded NMR spectra for each compounds can be found in supplementary material S2 and UV-VIS spectra can be found in S3.

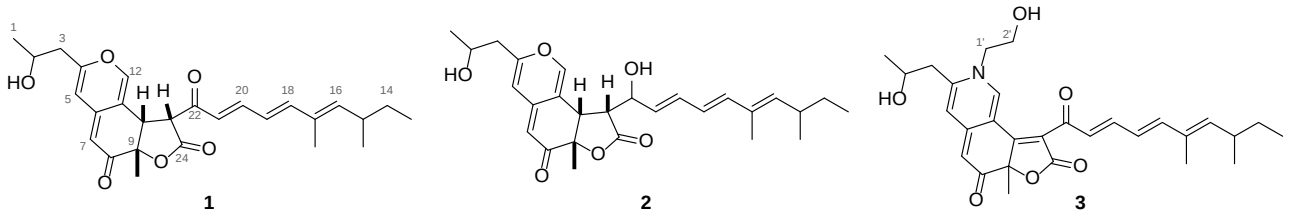


*Figure 3. Numbered structures of sassafrin E (**1**), sassafrin F (**2**), and sassafrinamine A (**3**).Figure 2*

In addition to the NMR experiments, we were able to further confirm the structure of sassafrin E, by exchanging H-23 with deuterium, and using the isotope labelled fragments in tandem MS experiments. In this way a proposed fragmentation pathways for the molecule could be suggested (spectra in supplementary material S4). What appears to be three separate fragmentation pathways is proposed based on the observed fragments (Figure 3).

[1] Correlations to C-6, C-10 and C-23 were not observed for **3**, and the tricyclic part of the structure was determined based on similarities with the remaining signals, compared to compound **1** and **2**.

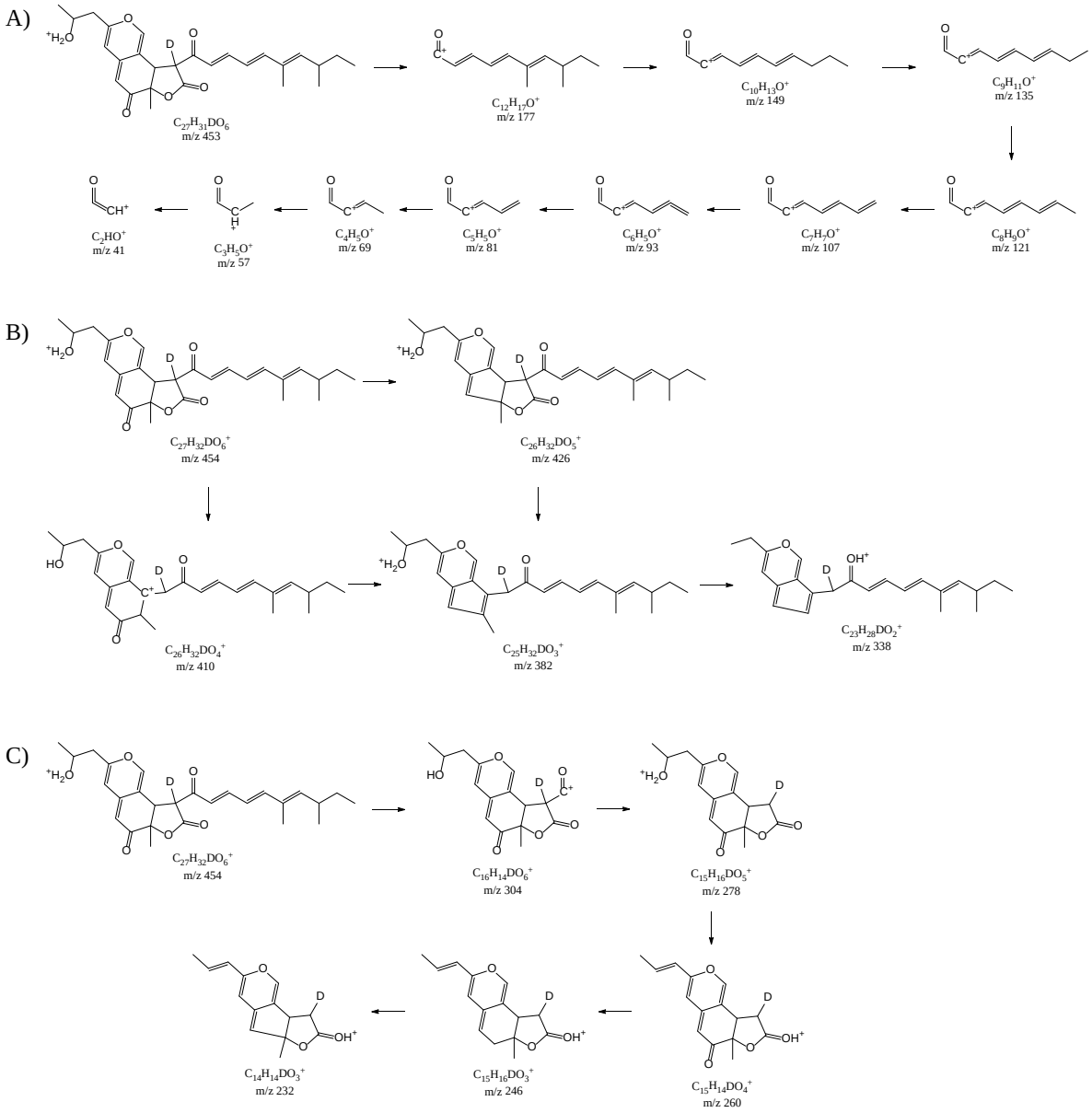


*Figure 4. Proposed MS/MS fragmentation pathways for sassafrin E (**1**), based on deuteration in position H-23.*

In summary, the isolated compounds are members of the compound class azaphilones, a diverse group of compounds, such as the ones obtained from various *Monascus* species, the so-called *Monascus* pigments.[10] The exact biosynthetic pathway of these compounds of have not been elucidated, though several proposals have been made, and it is expected that the compounds described in this study are synthesised in a similar fashion, as outlined in Figure 5.[11–16] However, in contrast to most other azaphilone pigments which are made from a hexaketide and a 3-oxo-fatty acid, sassafrin E and F, and sassafrinamine A are expected to be constructed from two polyketides, due to the relatively low level reduction of the second polyketide chain (C-13 to C-24). For Sassafrin F, we further expect the reduction at the ketone at C-22 to happen after construction of the compound backbone, as PKSs are usually stringent with regard to the reduction pattern

of their products, and it is unlikely that two different PKs would be incorporated into the same biosynthetic pathway. The fusion of the lactone ring also differ from *Monascus* pigments and is more similar to compounds such as chaetoviridins.[17] As we have recently discovered for the compound class atrorosins, the incorporation of nitrogen into the isochromene system can be done using various primary amine containing compounds.[18,19] We speculate that the nitrogen, and additional carbons and oxygen in sassafrinamine A originates from a decarboxylated serine, i.e. ethanolamine which is abundant in cells as constituents in phospholipids in cell membranes[20].

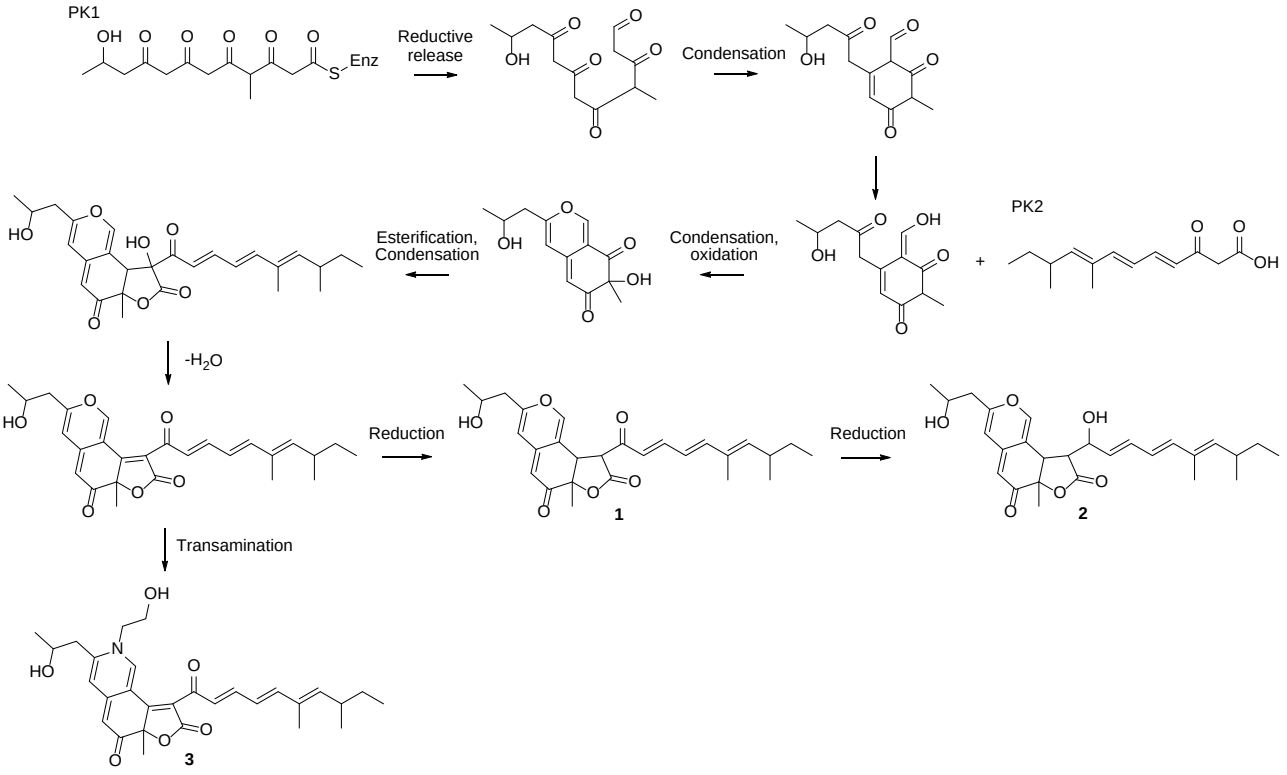


*Figure 5. Proposed biosynthetic pathway, for compounds **1**, **2** and **3**.*

Experimental

Solvents and instrumentation

All solvents were acquired from Sigma-Aldrich (St. Louis, Missouri, USA), ultra-pure water was made with a Milli-Q system (Millipore, Burlington, Massachusetts, USA).

Ultra-high Performance Liquid Chromatography-High Resolution Tandem Mass Spectrometry (UHPLC-HRMS/MS) was performed on an Agilent Infinity 1290 UHPLC system (Agilent Technologies, Santa Clara, CA, USA) equipped with a diode array detector. Separation was obtained on an Agilent Poroshell 120 phenyl-hexyl column (2.1 × 250 mm, 2.7 µm) with a linear gradient consisting of water (A) and acetonitrile (B) both

buffered with 20 mM formic acid, starting at 10% B and increased to 100% in 15 min where it was held for 2 min, returned to 10% in 0.1 min and remaining for 3 min (0.35 mL/min, 60 °C). MS detection was performed in positive detection mode on an Agilent 6545 QTOF MS equipped with Agilent Dual Jet Stream electrospray ion source with a drying gas temperature of 250 °C, gas flow of 8 L/min, sheath gas temperature of 300 °C and flow of 12 L/min. Capillary voltage was set to 4000 V and nozzle voltage to 500 V. Mass spectra were recorded at 10, 20 and 40 eV as centroid data for *m*/*z* 85–1700 in MS mode and *m*/*z* 30–1700 in MS/MS mode, with an acquisition rate of 10 spectra/s. Lock mass solution in 70:30 methanol:water was infused in the second sprayer using an extra LC pump at a flow of 15 µL/min using a 1:100 splitter. The solution contained 1 µM tributylamine (Sigma-Aldrich) and 10 µM Hexakis(2,2,3,3-tetrafluoropropoxy)phosphazene (Apollo Scientific Ltd., Cheshire, UK) as lock masses. The $[M + H]^+$ ions (*m*/*z* 186.2216 and 922.0098 respectively) of both compounds were used.

1D and 2D NMR spectra were recorded on a Bruker Avance 600 MHz or Bruker Avance 800 MHz spectrometer (Bruker, Billerica, MA, USA) located at the Department of Chemistry at the Technical University of Denmark. NMR spectra were acquired using standard pulse sequences. The solvent used was MeOD, which was also used as references with signals at δ_H = 3.31 ppm and δ_C = 49.0 ppm. Data processing and analysis was done using TopSpin 3.5pl7 (Bruker). *J*-couplings are reported in hertz (Hz) and chemical shifts in ppm (δ). Atrorosin Q was measured in DMSO-d_6, and the remaining compounds were measured in MeOD.

Strain and purification

The strain used for this study was *Aspergillus neoglaber* IBT3020, obtained from the DTU strain collection. For large scale extractions, the fungus was grown in six 2L conical flasks each with 500 mL of semi-liquid YES medium.

Extraction was done twice on the biomass using ethyl acetate acidified with 1% formic acid. Initial fractionation of the extract was done on an Isolera One (Biotage) flash system using a diol column eluted stepwise with dichloromethane (DCM), DCM:EtOAc (1:1), EtOAc, EtOAc:MeOH (1:1), and MeOH. Final isolation of the pure compounds was done using a semi-preparative Waters 600 Controller with a 996 photodiode array detector (Waters, Milford, MA, USA) equipped with a Luna II C18 column (250 × 10 mm, 5 µm, Phenomenex), using a H_2O/acetonitrile gradient with 50 ppm TFA.

References

1. Newman, D. J. & Cragg, G. M. Natural Products as Sources of New Drugs from 1981 to 2014. *J. Nat. Prod.* **79,** 629–661 (2016).
2. May Zin, W. *et al.* New Cyclotetrapeptides and a New Diketopiperzine Derivative from the Marine

Sponge-Associated Fungus Neosartorya glabra KUFA 0702. *Mar. Drugs* **14,** 136 (2016).

3. Jayasuriya, H. *et al.* Discovery and antibacterial activity of glabramycin A–C from Neosartorya glabra by an antisense strategy. *J. Antibiot. (Tokyo).* **62,** 265–269 (2009).

4. Liu, W., Zhao, H., Li, R., Zheng, H. & Yu, Q. Polyketides and Meroterpenoids from *Neosartorya glabra*. *Helv. Chim. Acta* **98,** 515–519 (2015).

5. Li, Y. Structural revision of glabramycins B and C, antibiotics from the fungus Neosartorya glabra by DFT calculations of NMR chemical shifts and coupling constants. *RSC Adv.* **5,** 36858–36864 (2015).

6. Otoguro, K. *et al.* Arisugacins C and D, Novel Acetylcholinesterase Inhibitors and Their Related Novel Metabolites Produced by Penicillium sp. FO-4259-11. *J. Antibiot. (Tokyo).* **53,** 50–57 (2000).

7. Otoguro, K., Kuno, F. & Omura, S. Arisugacins, selective acetylcholinesterase inhibitors of microbial origin. *Pharmacol. Ther.* **76,** 45–54 (1997).

8. Quang, D. N. *et al.* Sassafrins A–D, new antimicrobial azaphilones from the fungus Creosphaeria sassafras. *Tetrahedron* **61,** 1743–1748 (2005).

9. Stierle, A. A. *et al.* Azaphilones from an Acid Mine Extremophile Strain of a Pleurostomophora sp. *J. Nat. Prod.* **78,** 2917–2923 (2015).

10. Gao, J., Yang, S. & Qin, J. Azaphilones: Chemistry and Biology. *Chem. Rev.* **113,** 4755–811 (2013).

11. Hajjaj, H. *et al.* Biosynthetic Pathway of Citrinin in the Filamentous Fungus Monascus ruber as Revealed by 13C Nuclear Magnetic Resonance. *Appl. Environ. Microbiol.* **65,** 311–314 (1999).

12. Woo, P. C. Y. *et al.* The biosynthetic pathway for a thousand-year-old natural food colorant and citrinin in Penicillium marneffei. *Sci. Rep.* **4,** 6728 (2014).

13. Liu, Q. *et al.* MpigE, a gene involved in pigment biosynthesis in Monascus ruber M7. *Appl. Microbiol. Biotechnol.* **98,** 285–296 (2014).

14. Tolborg, G., Isbrandt, T., Larsen, T. O. & Workman, M. in *Bio-pigmentation and Biotechnological Implementations* (ed. Singh, O. V.) 23–51 (John Wilety & Sons, Inc., 2017).

15. Chen, W. *et al.* Orange, red, yellow: biosynthesis of azaphilone pigments in Monascus fungi. *Chem. Sci.* **8,** 4917–4925 (2017).

16. Somoza, A. D., Lee, K.-H., Chiang, Y.-M., Oakley, B. R. & Wang, C. C. C. Reengineering an Azaphilone Biosynthesis Pathway in Aspergillus nidulans To Create Lipoxygenase Inhibitors. *Org. Lett.* **14,** 972–975 (2012).

17. Winter, J. M. *et al.* Identification and Characterization of the Chaetoviridin and Chaetomugilin Gene Cluster in Chaetomium globosum Reveal Dual Functions of an Iterative Highly-Reducing Polyketide Synthase. *J. Am. Chem. Soc.* **134,** 17900–17903 (2012).

18. Isbrandt, T., Tolborg, G., Workman, M. & Larsen, T. O. IN PREP: Atrorosins: a new subgroup of Monascus pigments from Talaromyces atroroseus.

19. Tolborg, G., Isbrandt, T., Ødum, A., Larsen, T. O. & Workman, M. IN PREP: Unique processes yielding pure azaphilones in Talaromyces atroroseus.

20. Wellner, N., Diep, T. A., Janfelt, C. & Hansen, H. S. N-acylation of phosphatidylethanolamine and its biological functions in mammals. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1831,** 652–662 (2013).

Supplementary material

New azaphilones from *Aspergillus neoglaber*

Thomas Isbrandt and Thomas Ostenfeld Larsen

Department of Biotechnology and Biomedicine, Technical University of Denmark, Kongens Lyngby, Denmark

Contents:

S1. Dereplication of mz453 compound in *Aspergillus neoglaber* compared to *Penicillium echinulatum*.

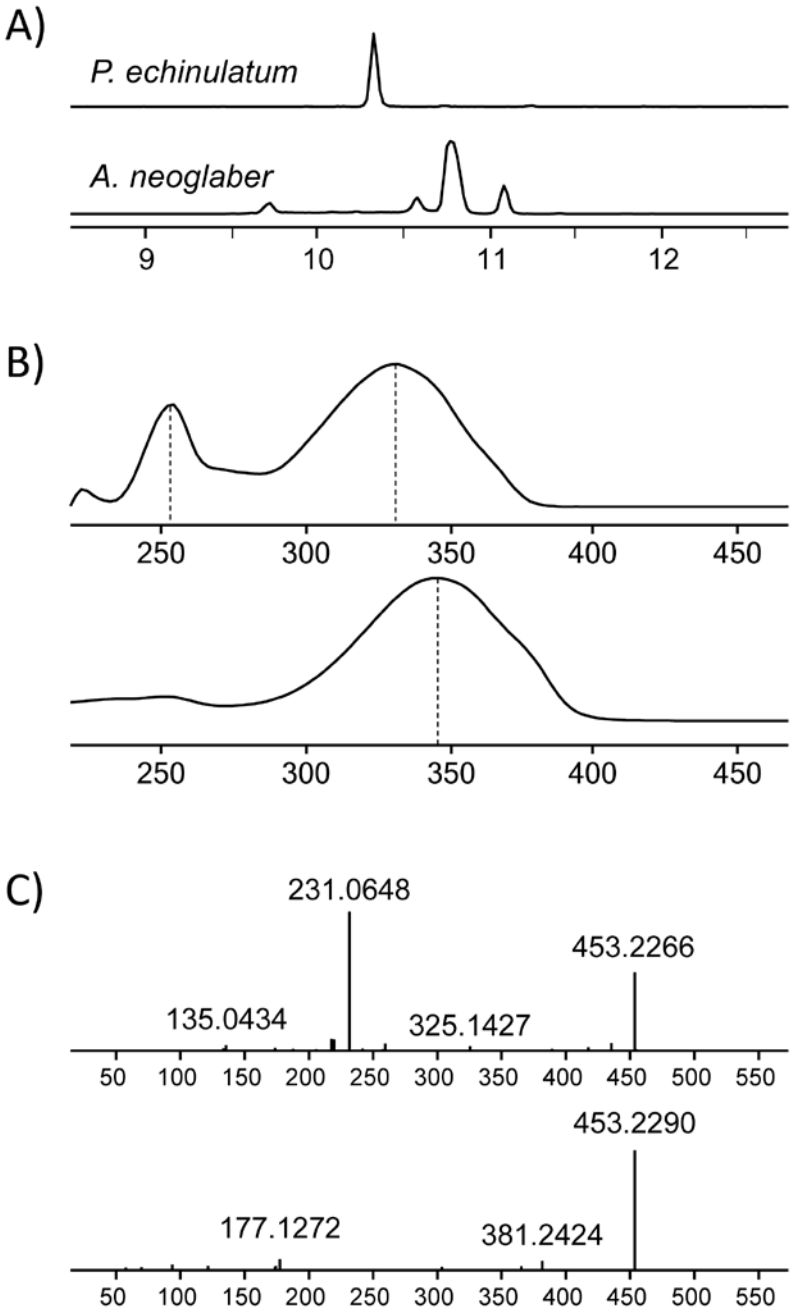
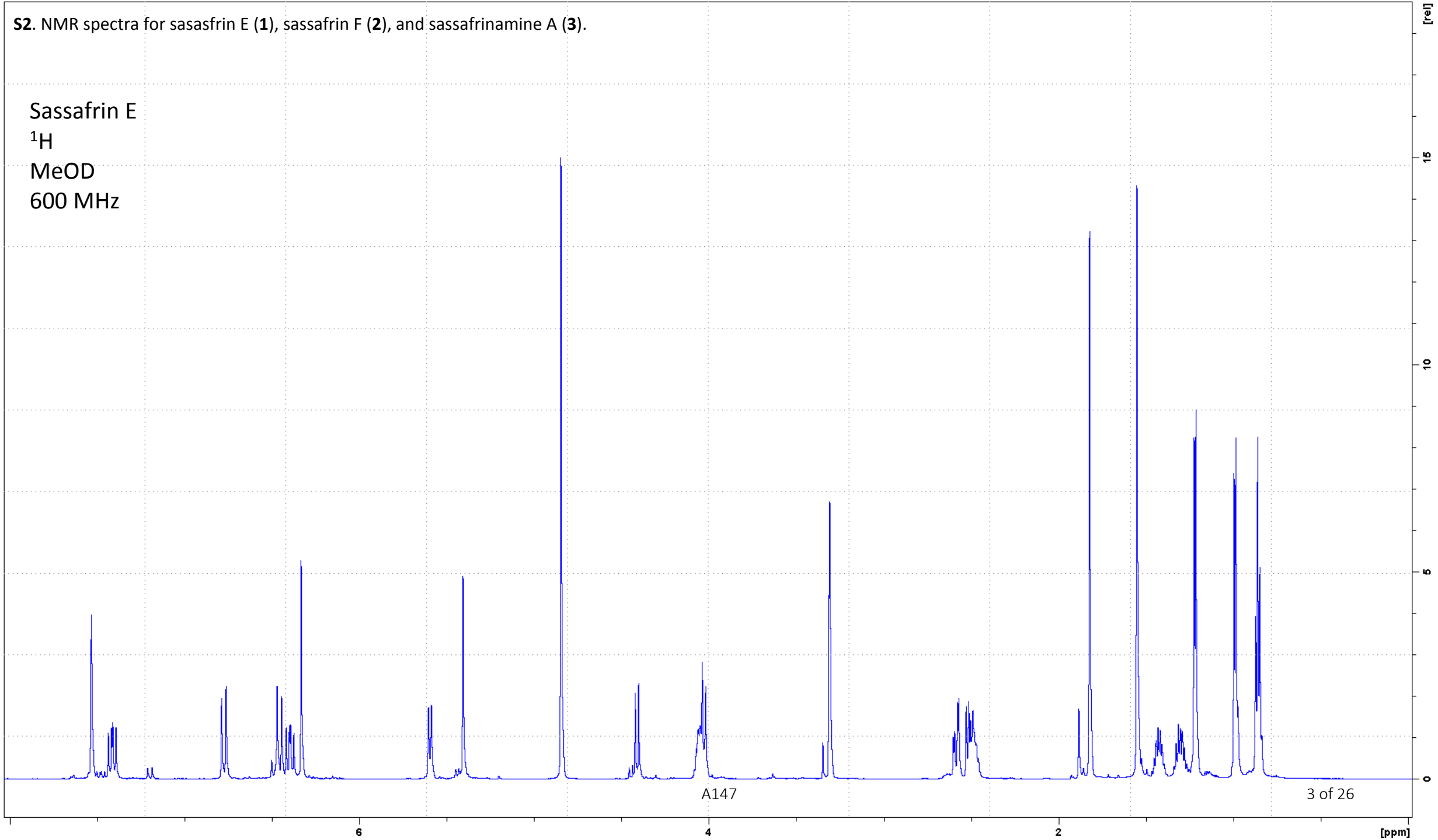


Figure S1. A) Extracted ion chromatogram (m/z 453.2277) for extracts from P. echinulatum and A. neoglaber. B) UV-VIS spectra for the mz453 compounds in P. echinulatum (top) and A. neoglaber (bottom). C) Tandem mass spectra (20 eV) for the mz453 compounds in P. echinulatum (top) and A. neoglaber (bottom).

S2. NMR spectra for sasasfrin E (**1**), sassafrin F (**2**), and sassafrinamine A (**3**).

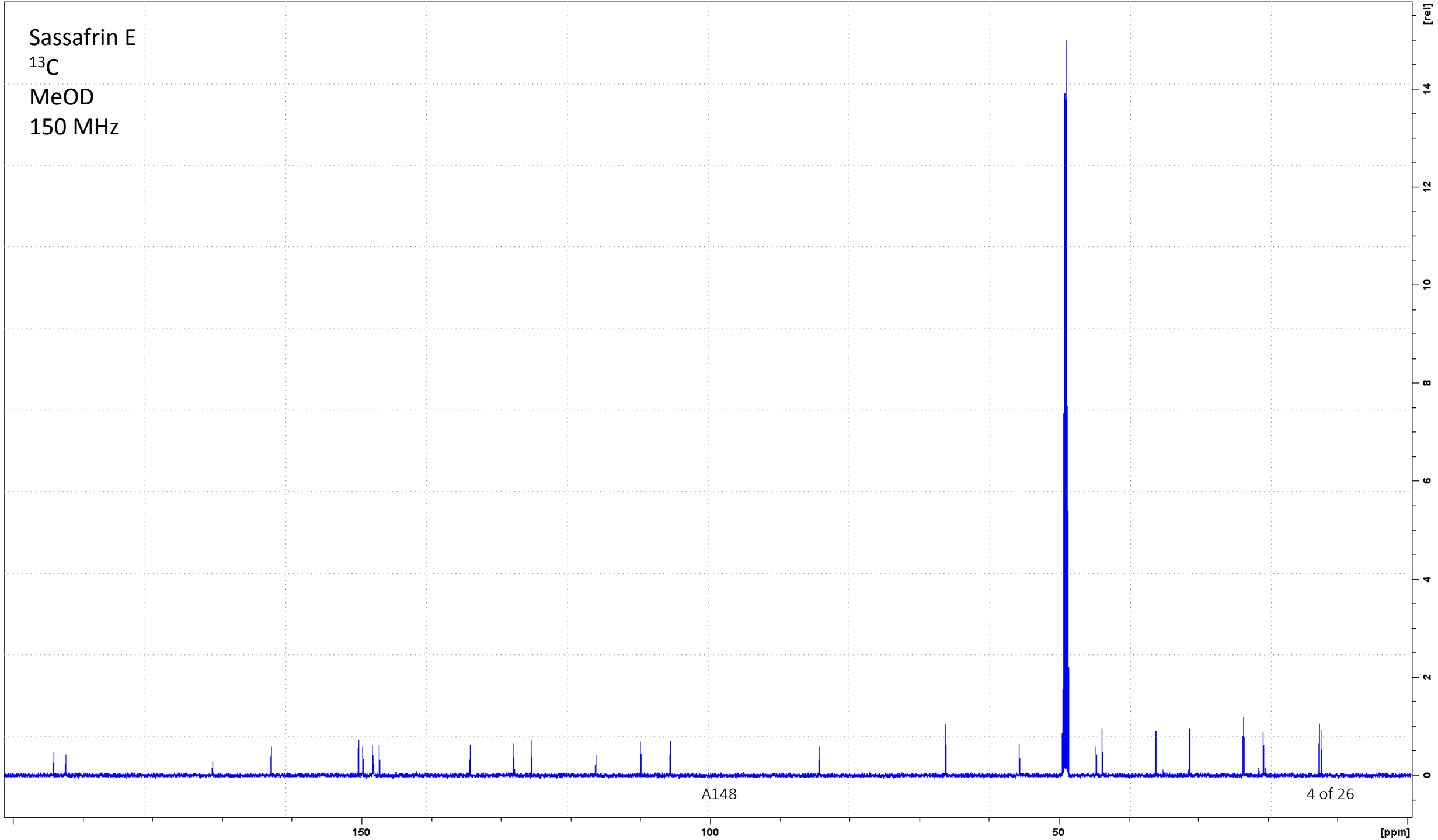


Sassafrin E

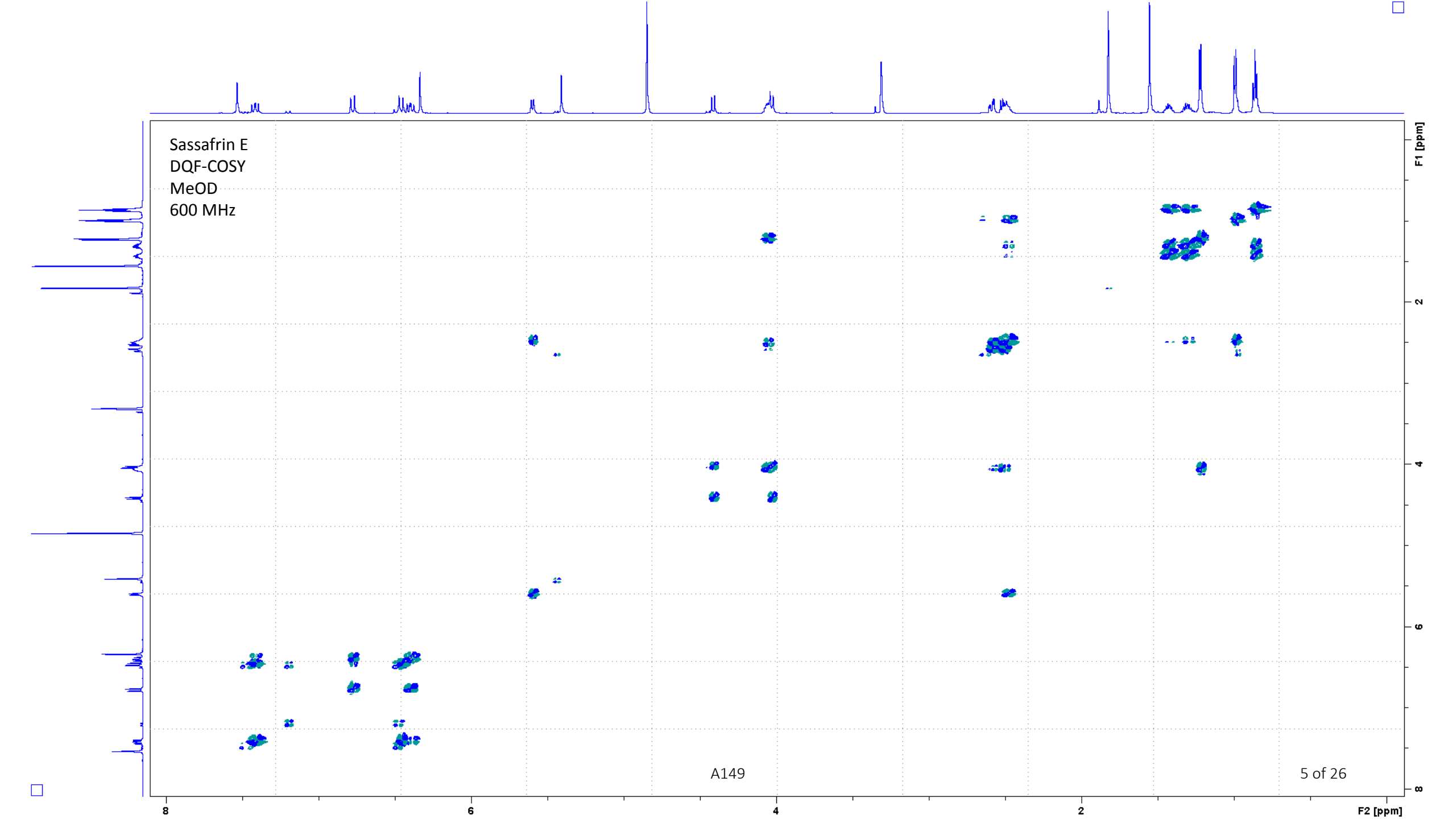
^{13}C

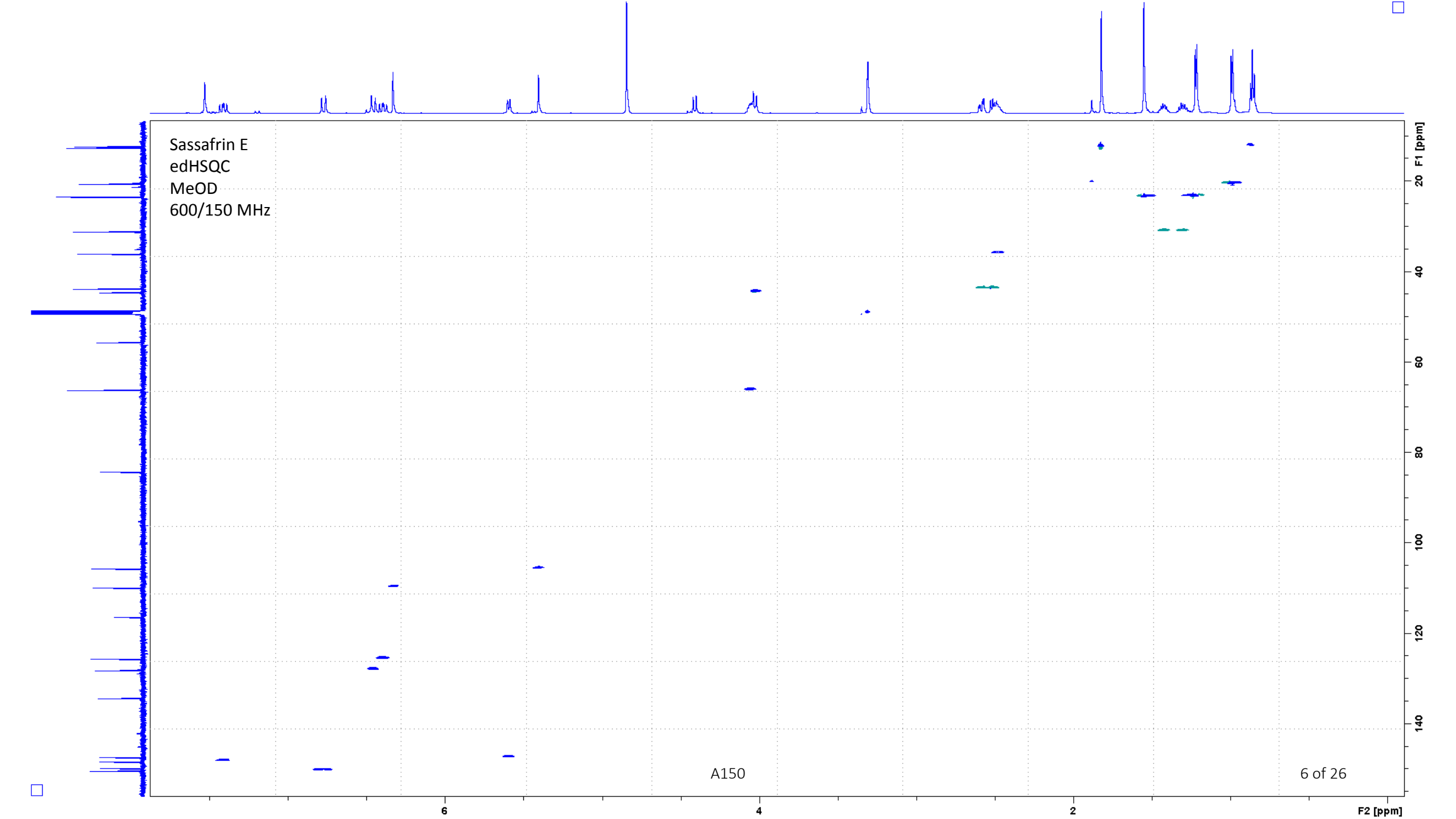
MeOD

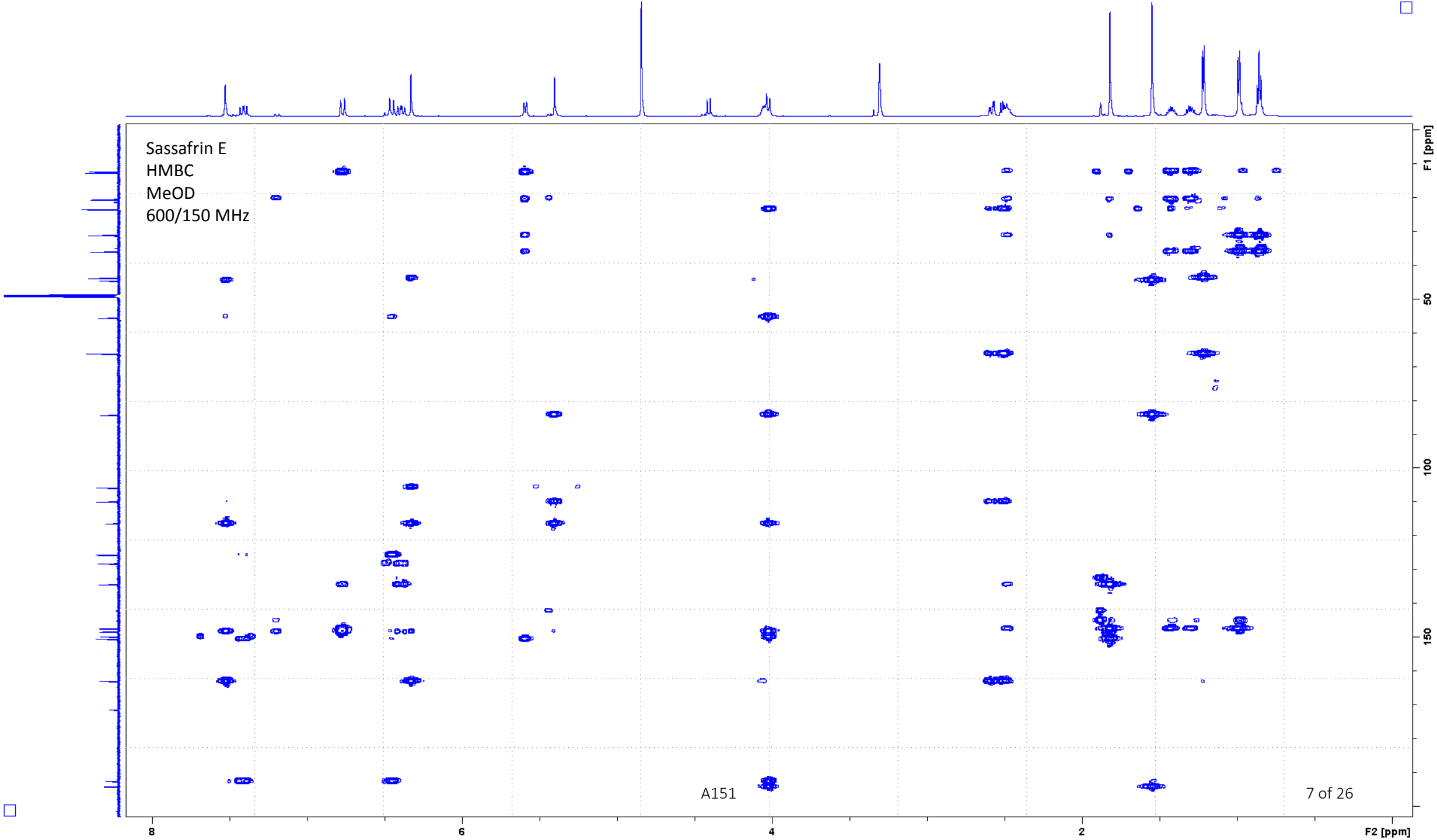
150 MHz



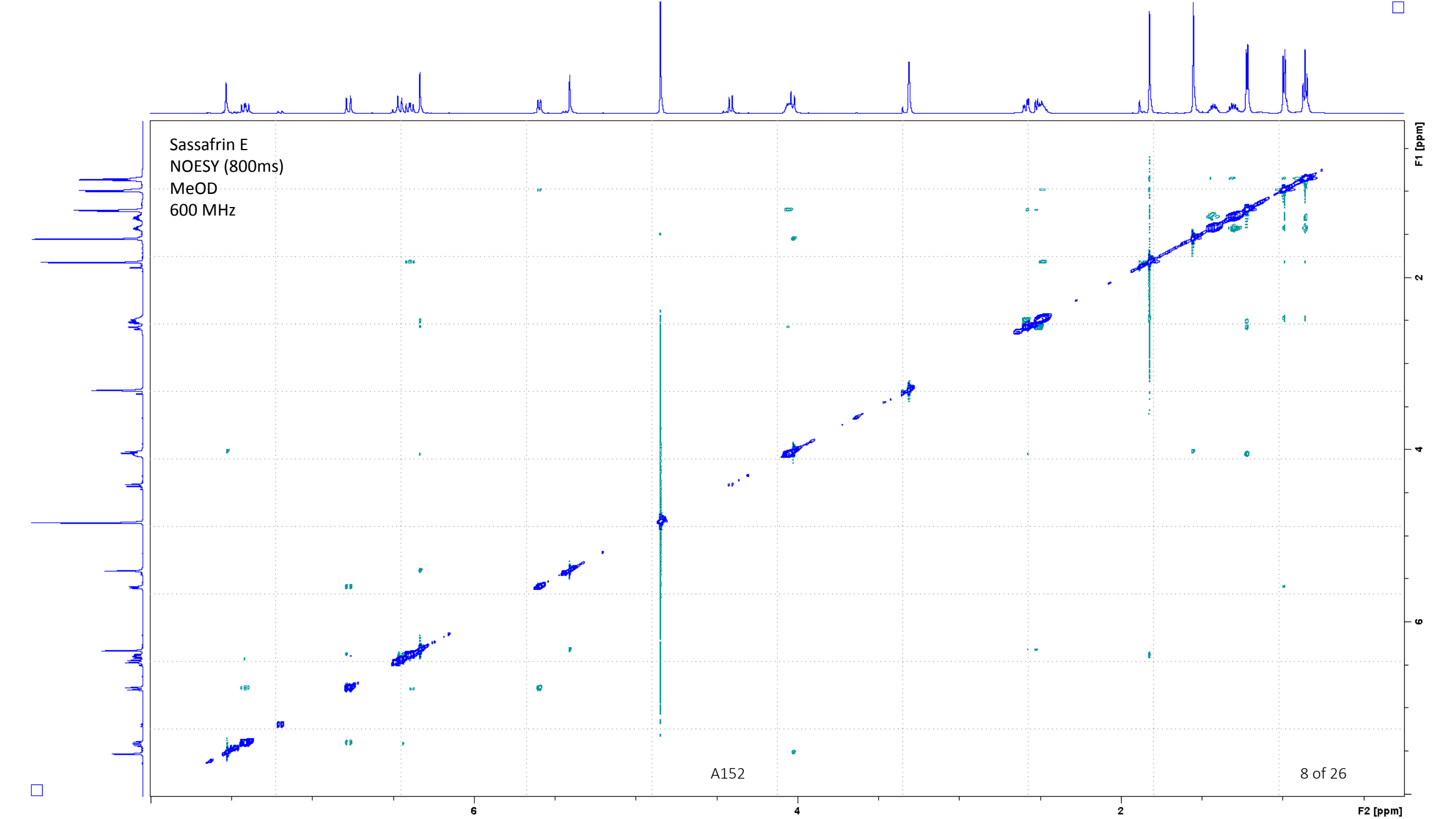
A148

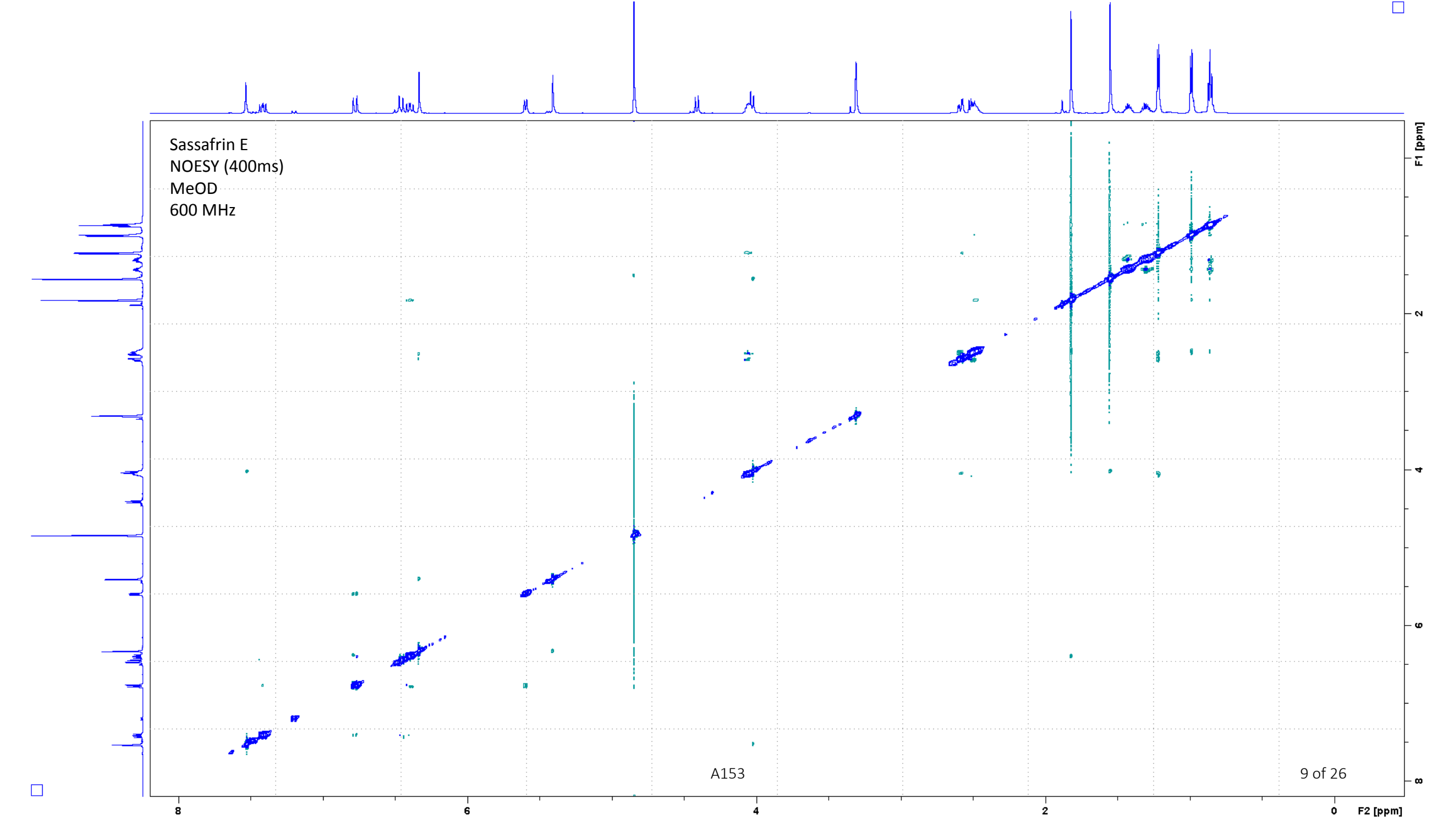

Sassafrin E
DQF-COSY
MeOD
600 MHz
F1 [ppm]
F2 [ppm]
8
6
4
2
A149
5 of 26

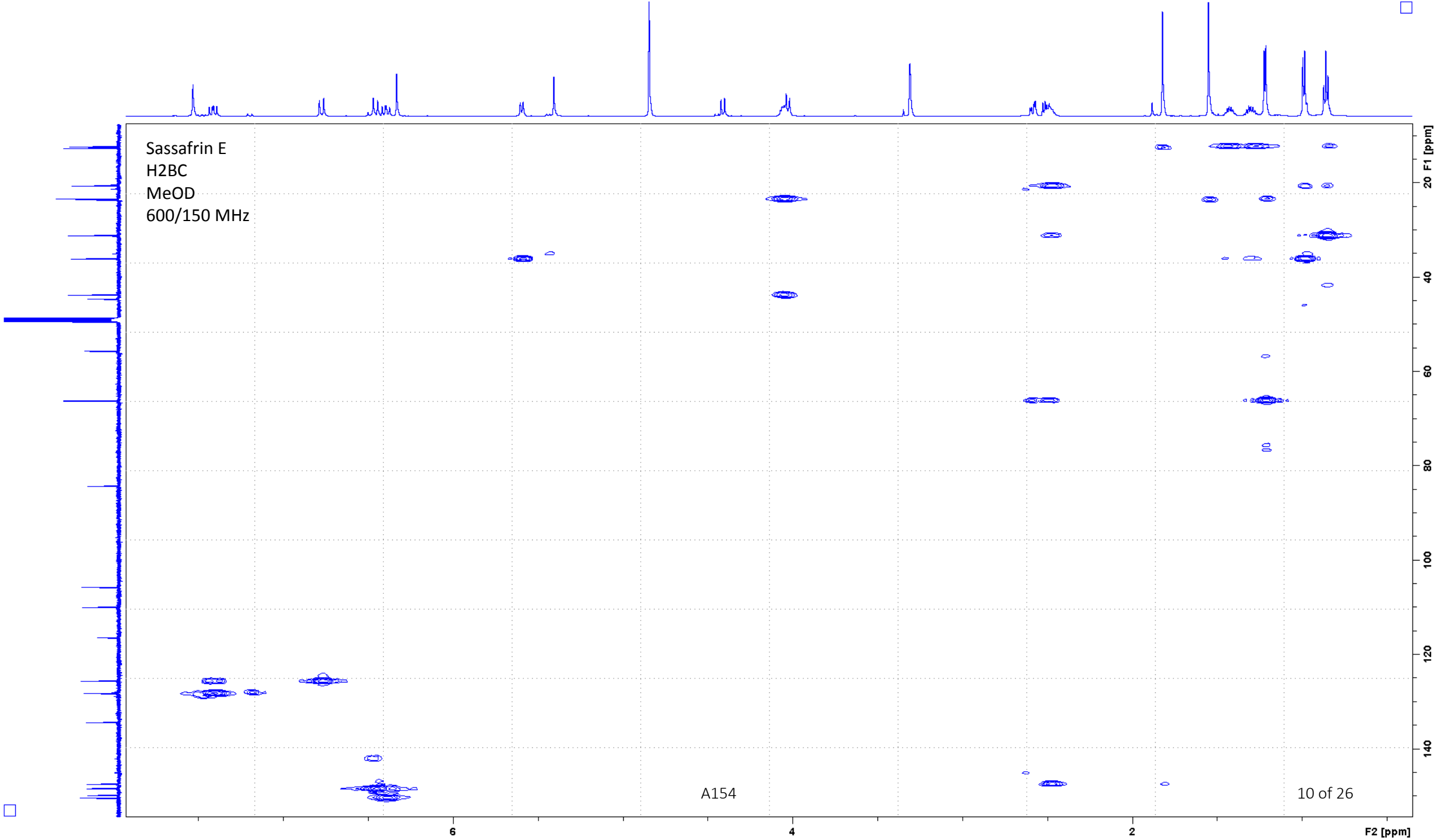

Sassafrin E
edHSQC
MeOD
600/150 MHz
A150
6 of 26
F1 [ppm]
F2 [ppm]
20
40
60
80
100
120
140
6
4
2

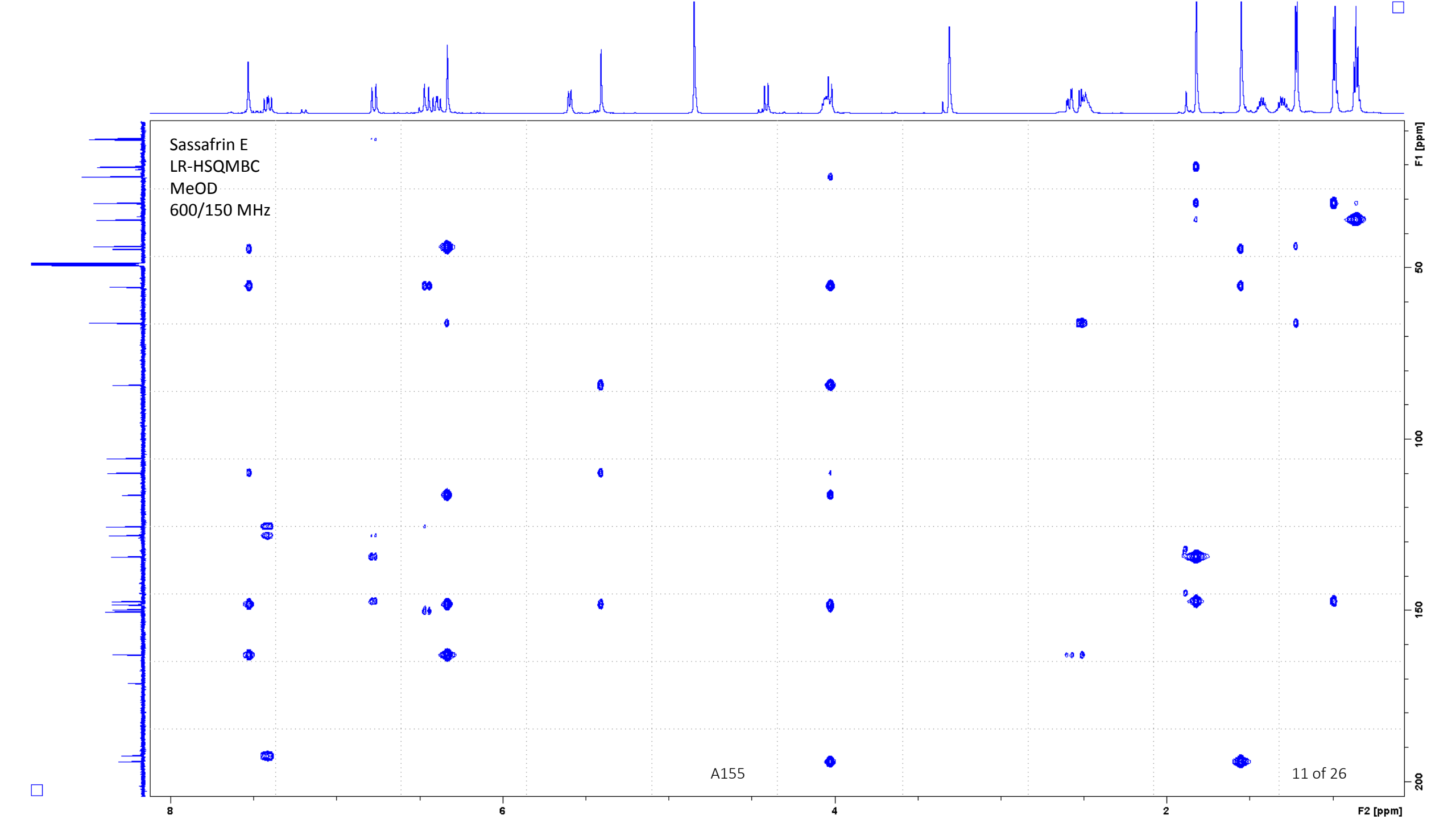


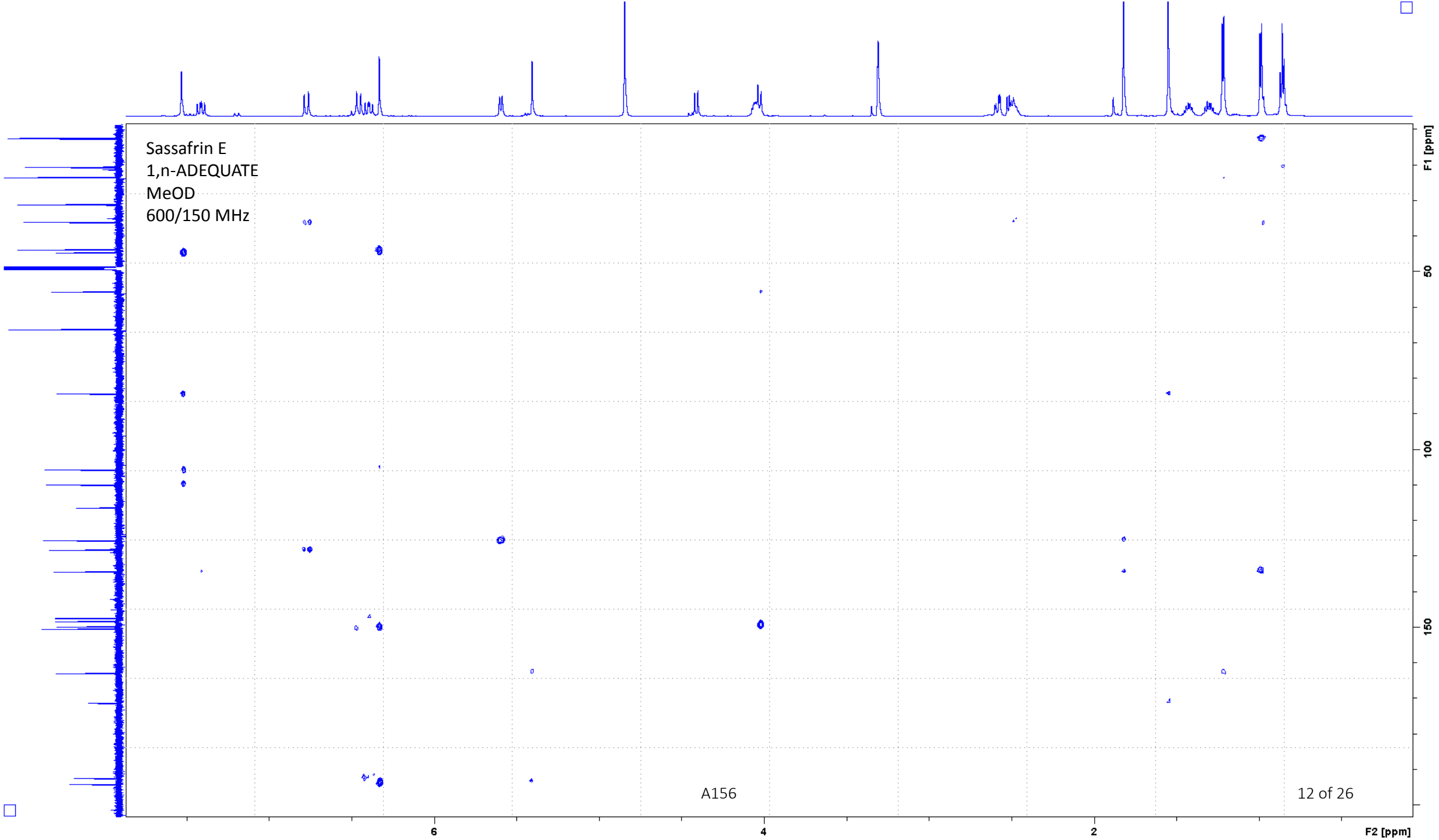
A151

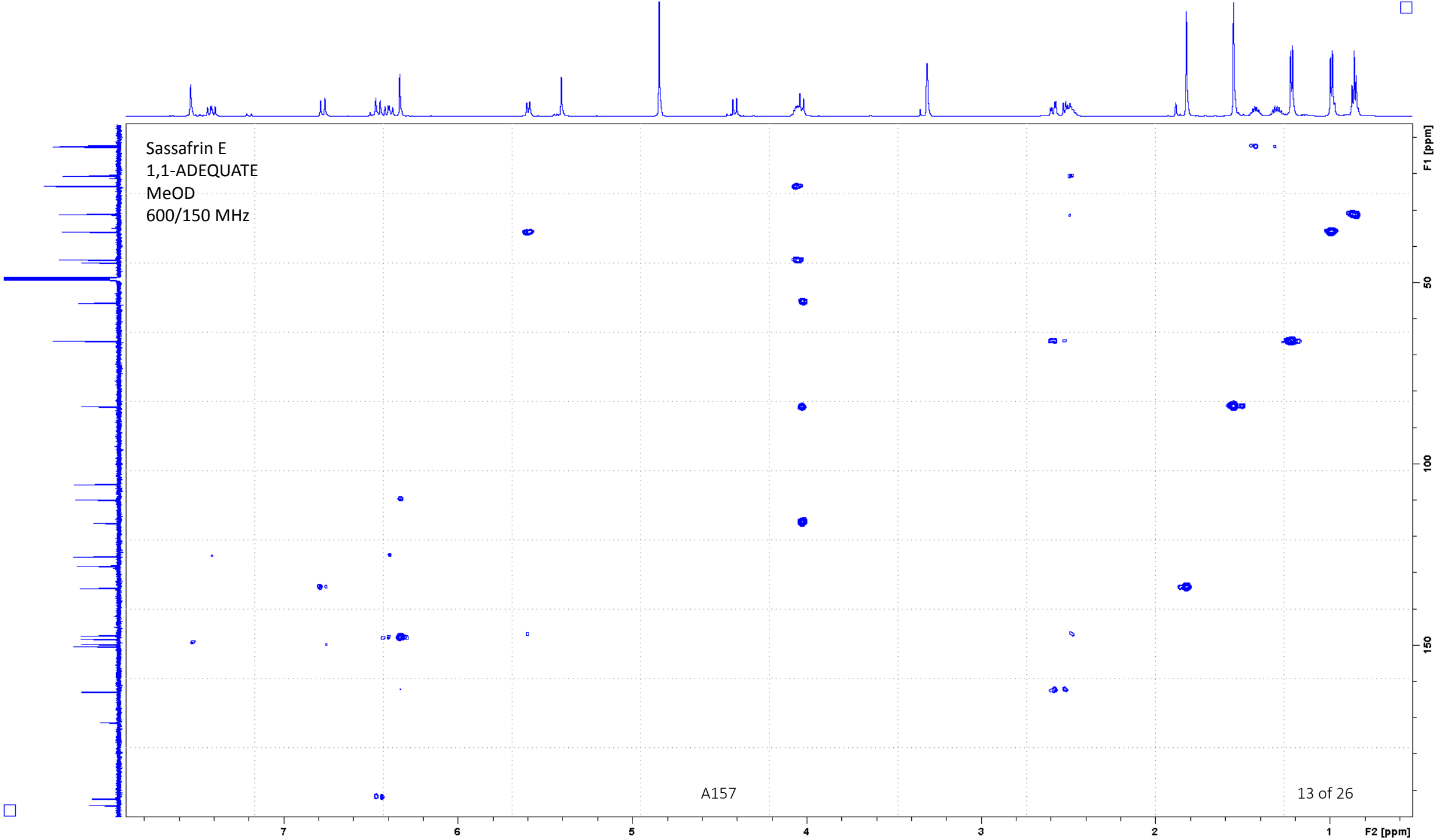

Sassafrin E
NOESY (800ms)
MeOD
600 MHz
A152
8 of 26
F1 [ppm]
F2 [ppm]
2
4
6

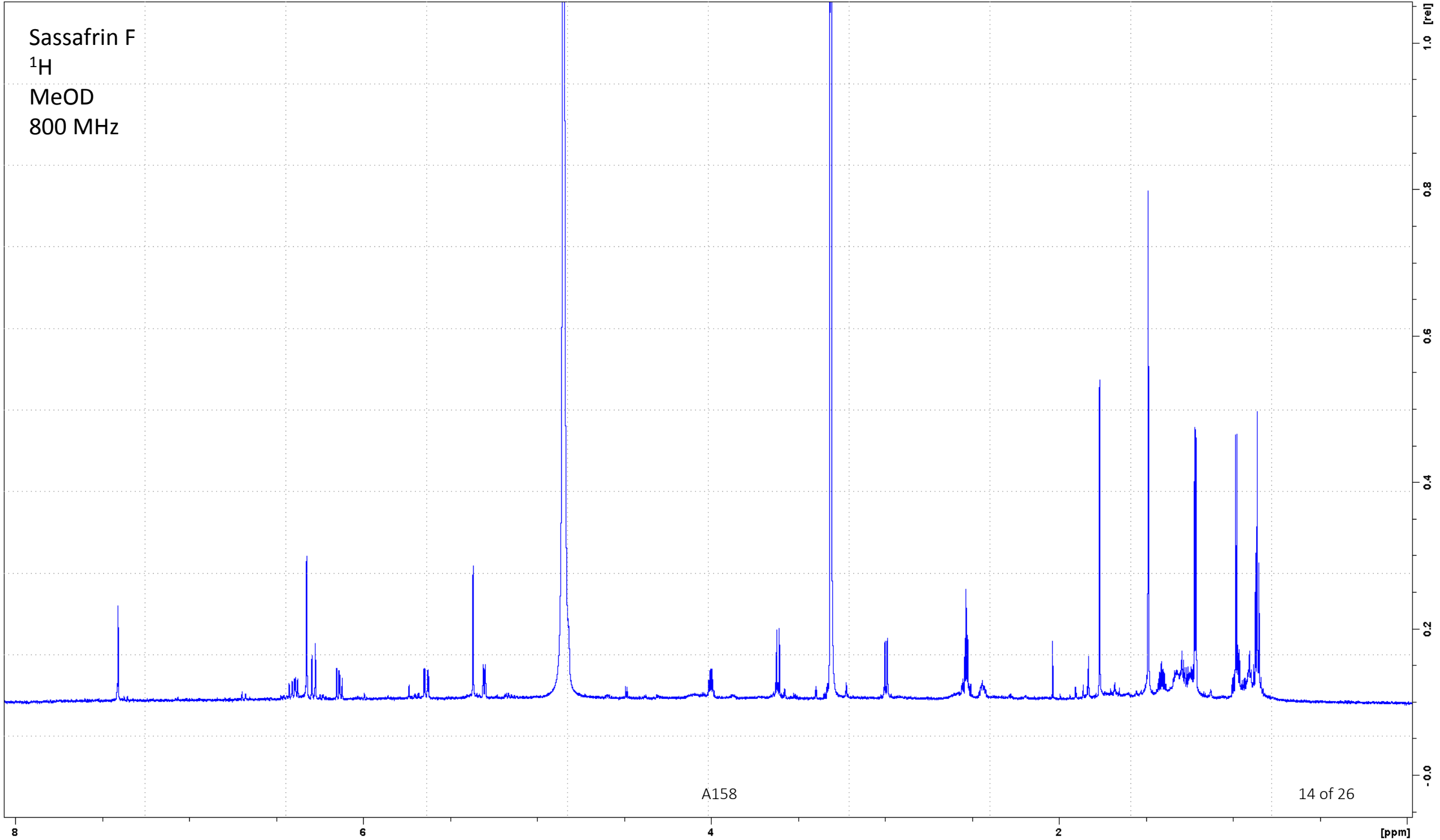

Sassafrin E
NOESY (400ms)
MeOD
600 MHz
A153
9 of 26
F1 [ppm]
F2 [ppm]
8
6
4
2
0
2
4
6
8

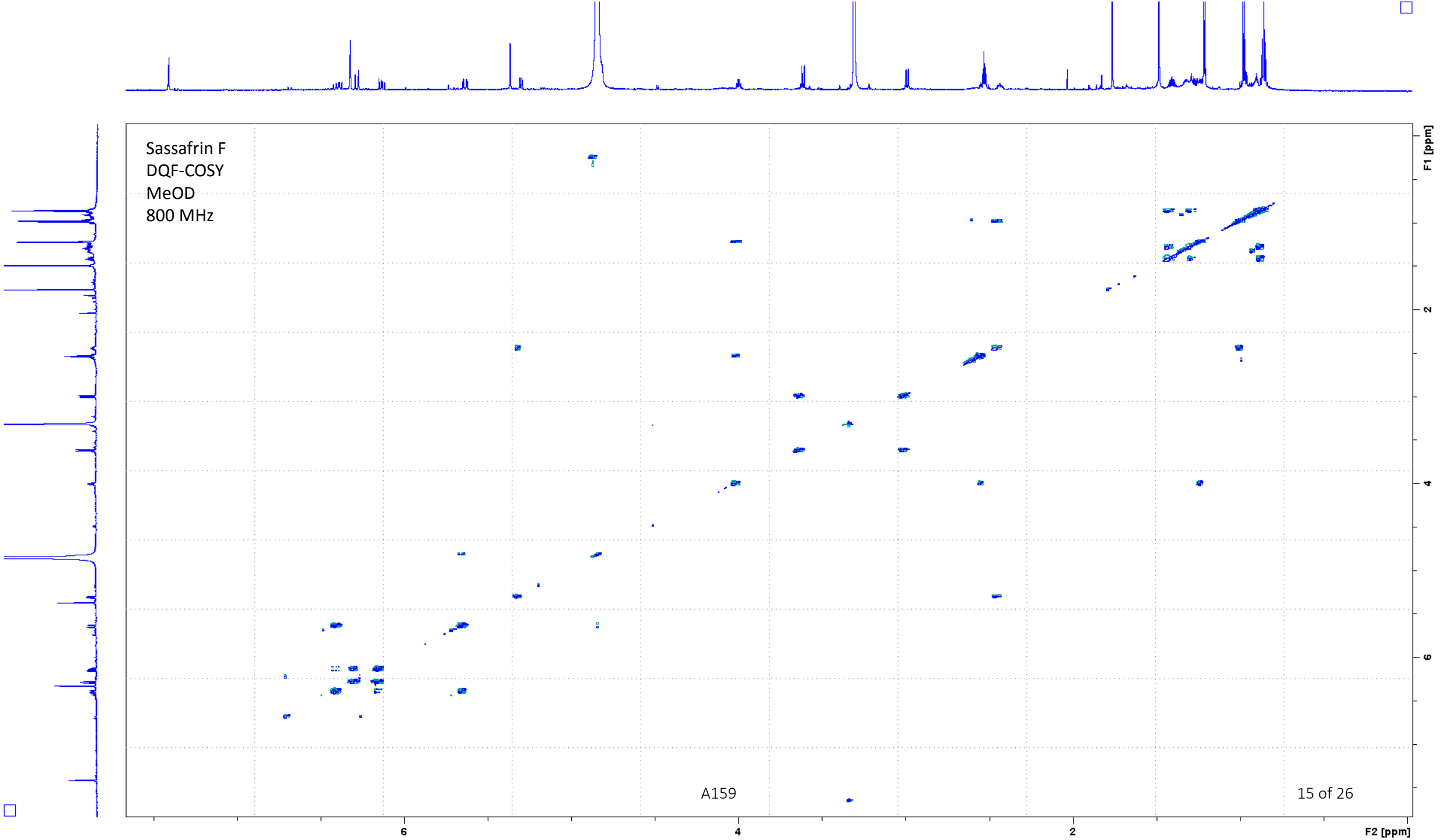

Sassafrin E
H2BC
MeOD
600/150 MHz
A154
10 of 26
F1 [ppm]
F2 [ppm]
20
40
60
80
100
120
140
6
4
2

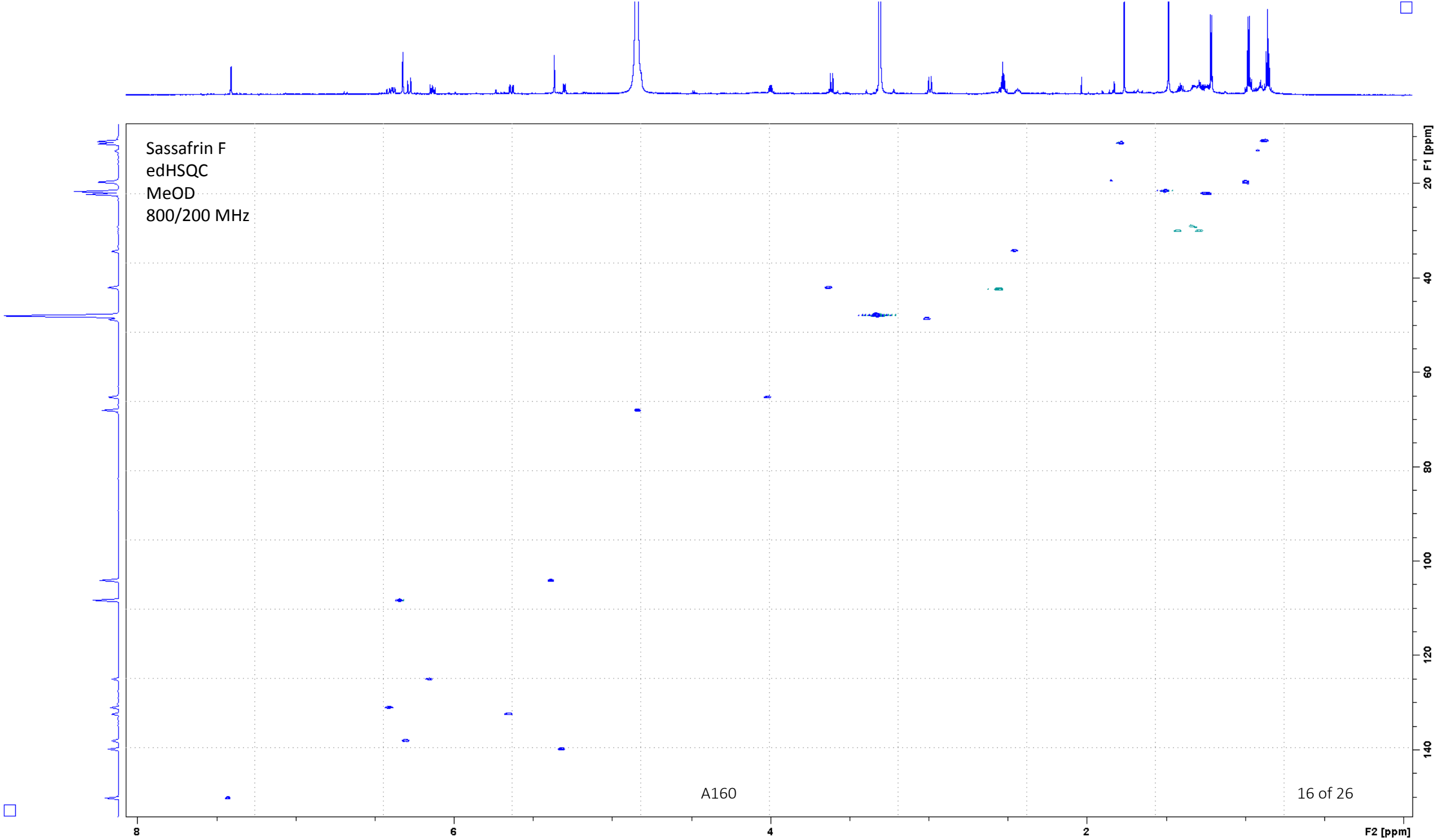

Sassafrin E
LR-HSQMBC
MeOD
600/150 MHz
A155
11 of 26
F1 [ppm]
F2 [ppm]
50
100
150
200
8
6
4
2

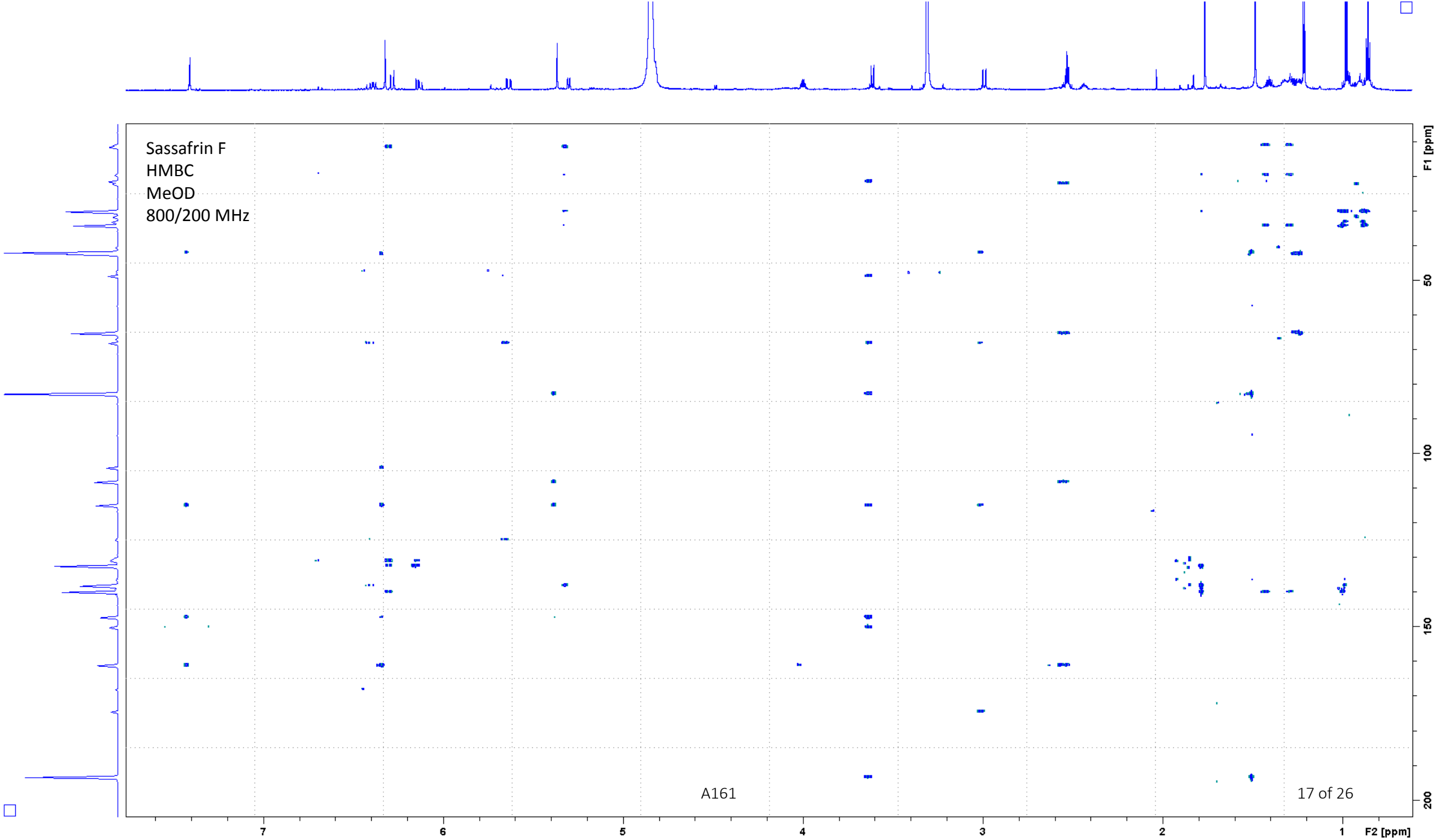

Sassafrin E
1,n-ADEQUATE
MeOD
600/150 MHz
F1 [ppm]
50
100
150
6
4
2
F2 [ppm]

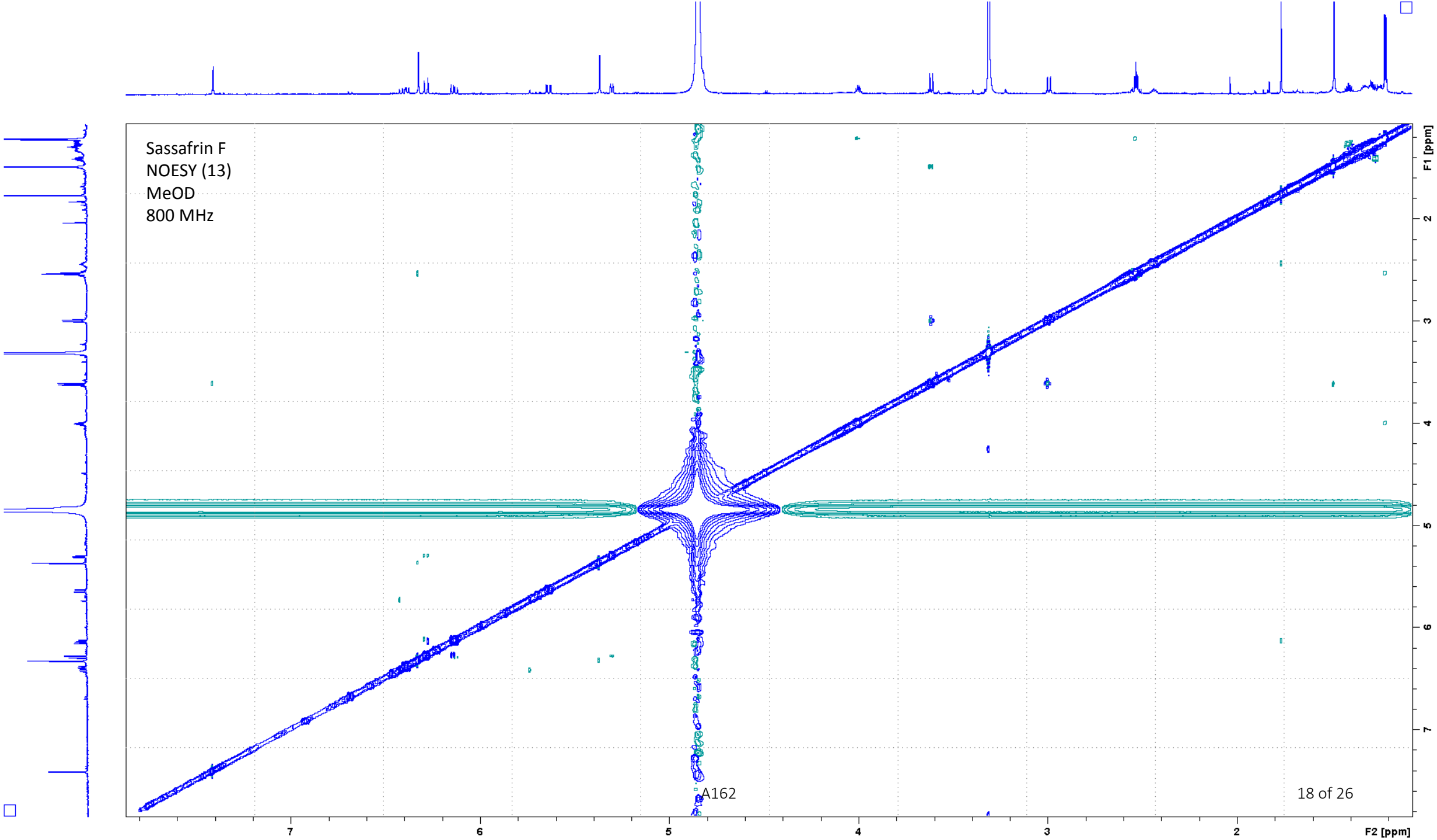

Sassafrin E
1,1-ADEQUATE
MeOD
600/150 MHz
A157
13 of 26
F1 [ppm]
F2 [ppm]
50
100
150
7
6
5
4
3
2
1

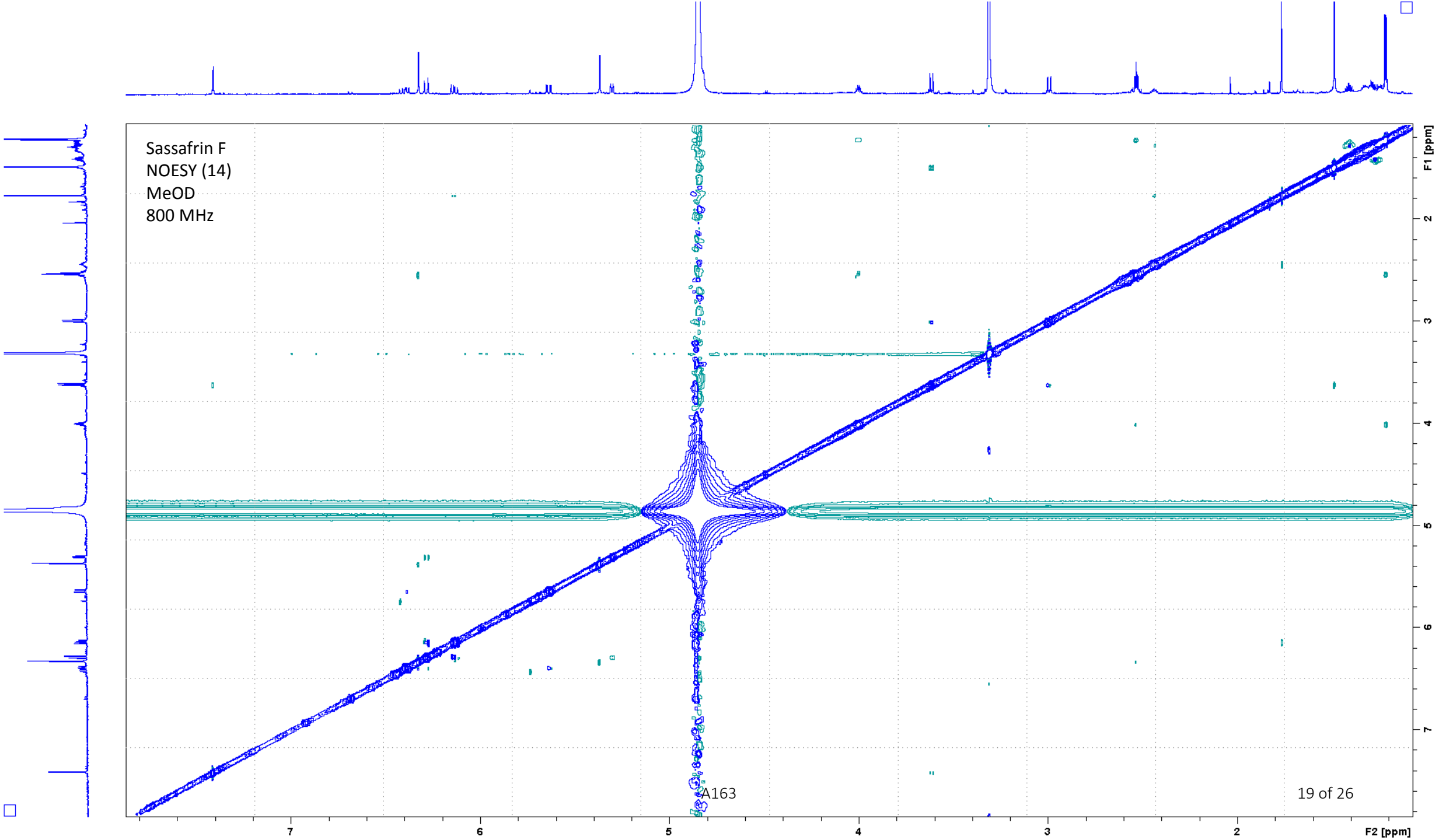

Sassafrin F
1H
MeOD
800 MHz
1.0 [rel]
0.8
0.6
0.4
0.2
-0.0
8
6
4
2
[ppm]

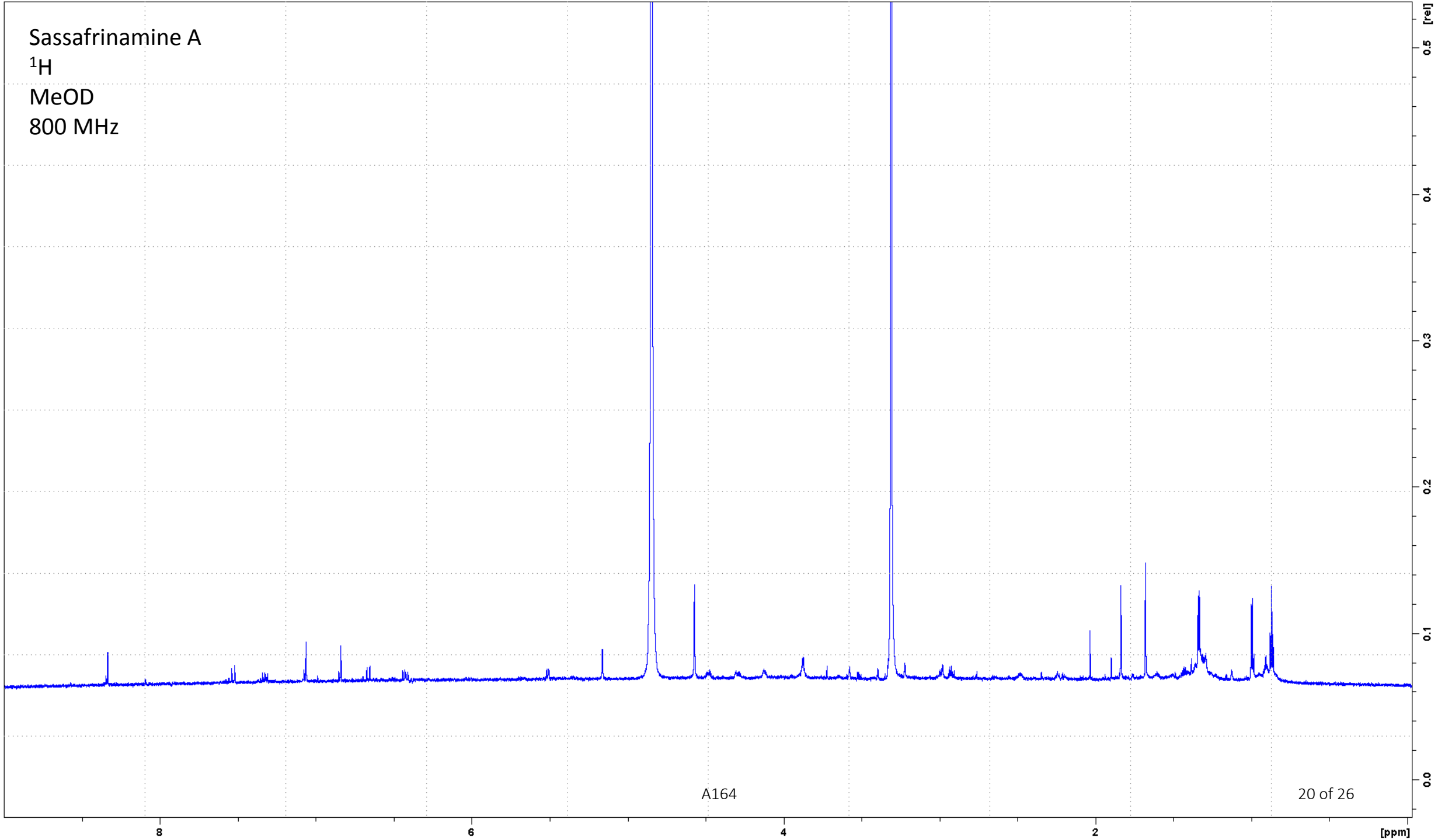

Sassafrin F
DQF-COSY
MeOD
800 MHz
A159
15 of 26
F1 [ppm]
F2 [ppm]
2
4
6
6
4
2

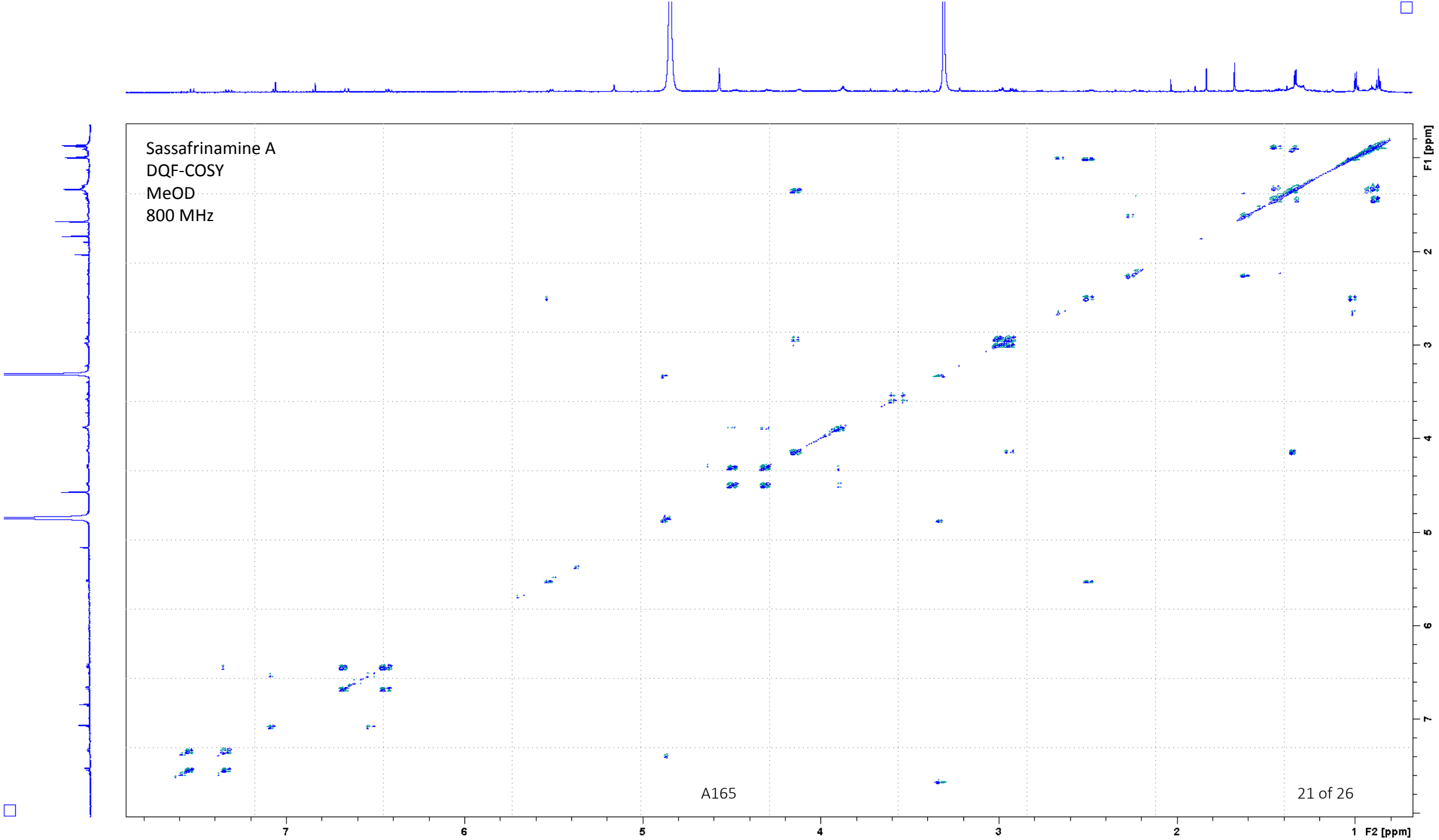

Sassafrin F
edHSQC
MeOD
800/200 MHz
A160
16 of 26
F1 [ppm]
F2 [ppm]
20
40
60
80
100
120
140
8
6
4
2

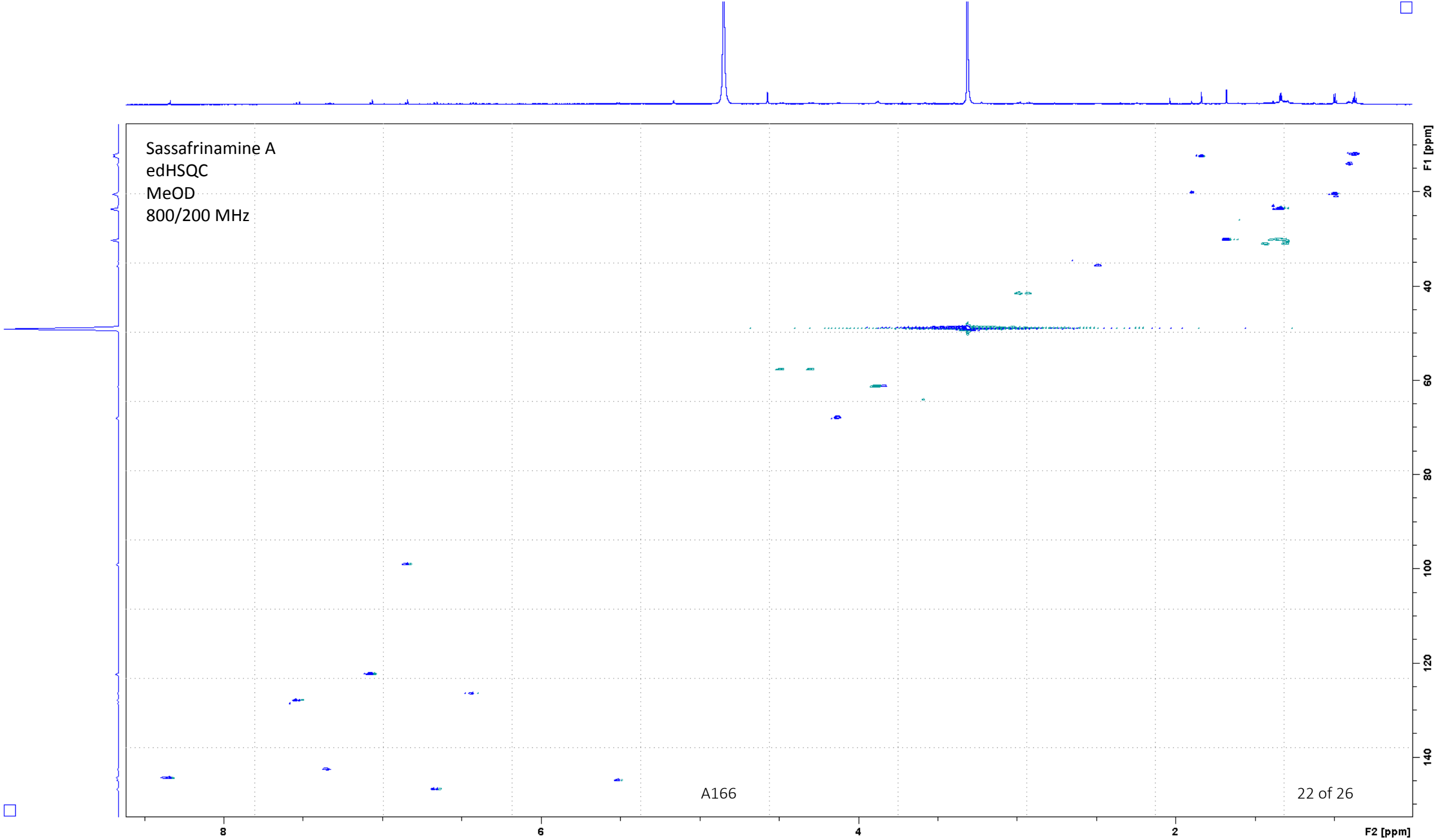

Sassafrin F
HMBC
MeOD
800/200 MHz
A161
17 of 26
F1 [ppm]
F2 [ppm]
50
100
150
200
7
6
5
4
3
2
1

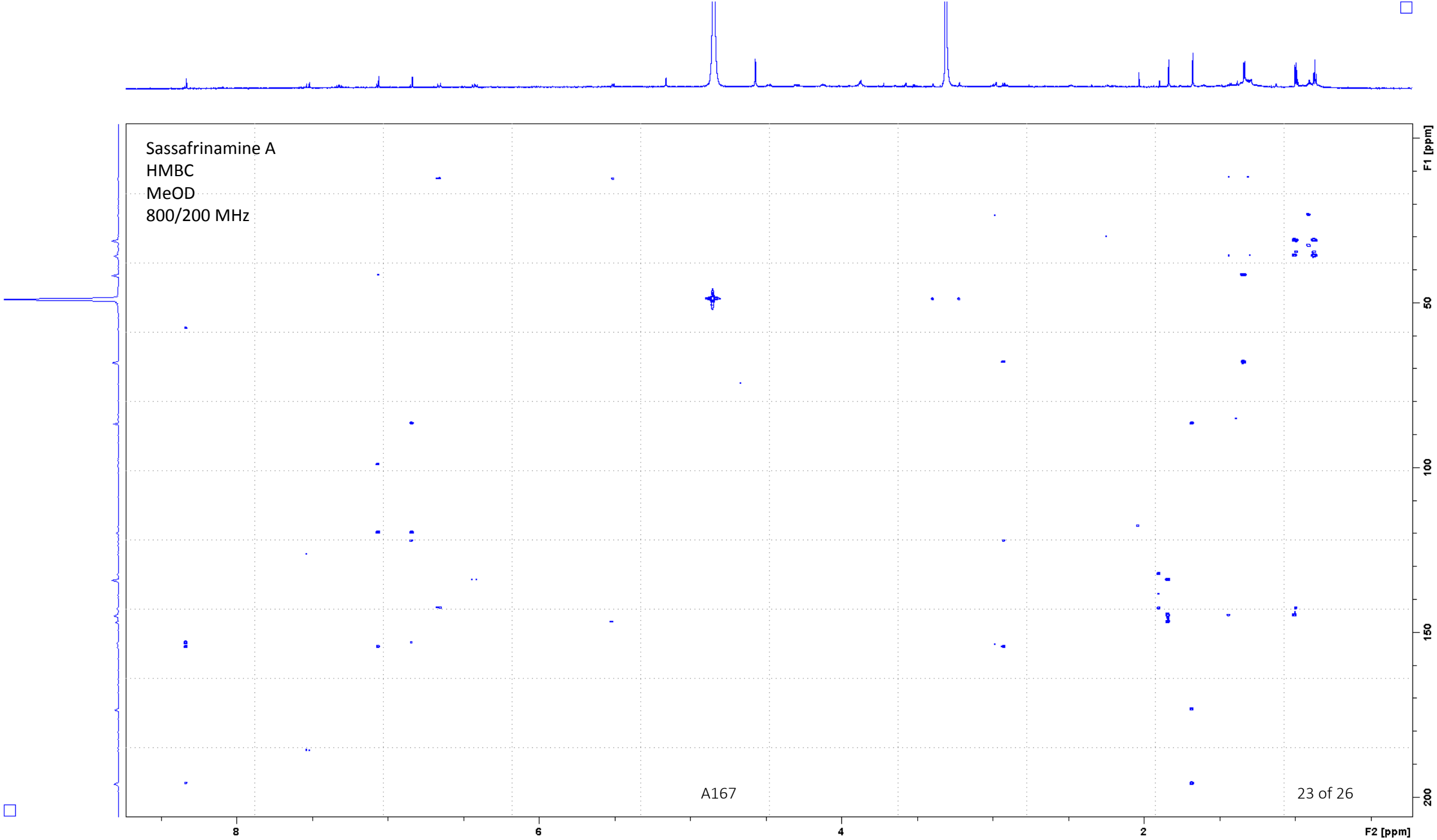

Sassafrin F
NOESY (13)
MeOD
800 MHz
F1 [ppm]
F2 [ppm]
7
6
5
4
3
2
A162


Sassafrin F
NOESY (14)
MeOD
800 MHz
F1 [ppm]
F2 [ppm]
7
6
5
4
3
2


Sassafrinamine A
1H
MeOD
800 MHz
A164
20 of 26
8
6
4
2
[ppm]
0.0
0.1
0.2
0.3
0.4
0.5
[rel]


Sassafrinamine A
DQF-COSY
MeOD
800 MHz
A165
21 of 26
F1 [ppm]
F2 [ppm]
1
2
3
4
5
6
7


Sassafrinamine A
edHSQC
MeOD
800/200 MHz
A166
22 of 26
F1 [ppm]
F2 [ppm]
20
40
60
80
100
120
140
8
6
4
2


Sassafrinamine A
HMBC
MeOD
800/200 MHz
A167
23 of 26
F1 [ppm]
F2 [ppm]
8
6
4
2
50
100
150
200

S3. UV-VIS spectra for A) sassafrin E, B) sassafrin F, and C) sassafrinamine A.

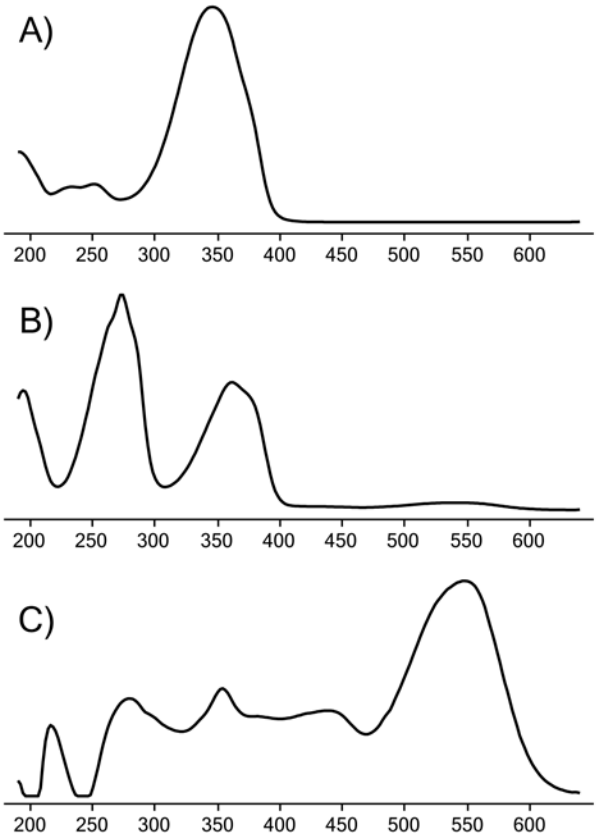


Figure S 4. UV-VIS spectra for A) Sassafrin E, B) Sassafrin F, C) Sassafrinamine A. All spectra were recorded during UHPLC analysis in acetonitrile and H_2O, acidified with 20mM fromic acid.

S4. Tandem mass spectra for deuterated sassafrin E (**1**).

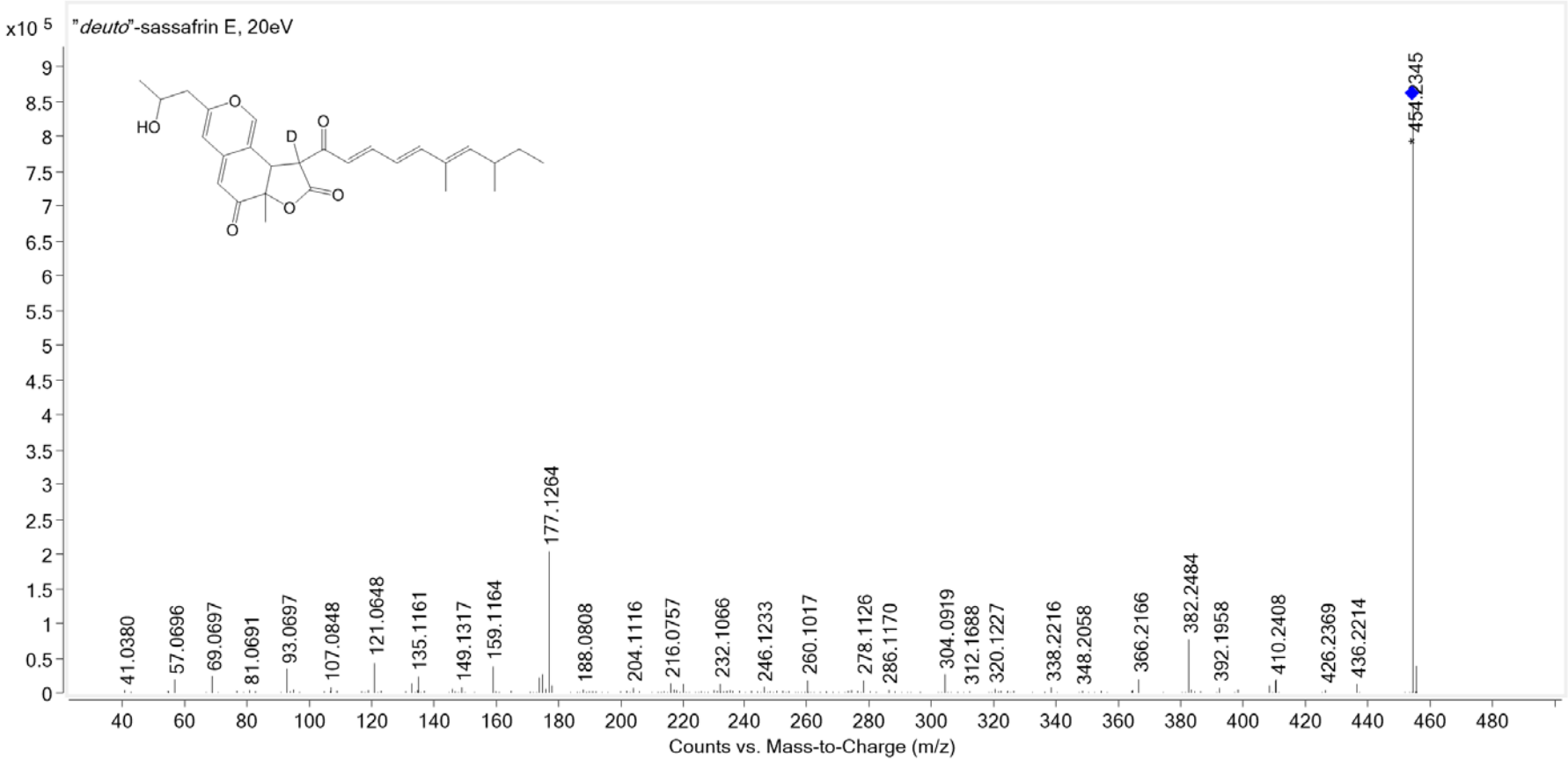


*Figure S2. Tandem mass spectrum for deuterated sassafrin E (**1**) at 20 eV colission energy.*

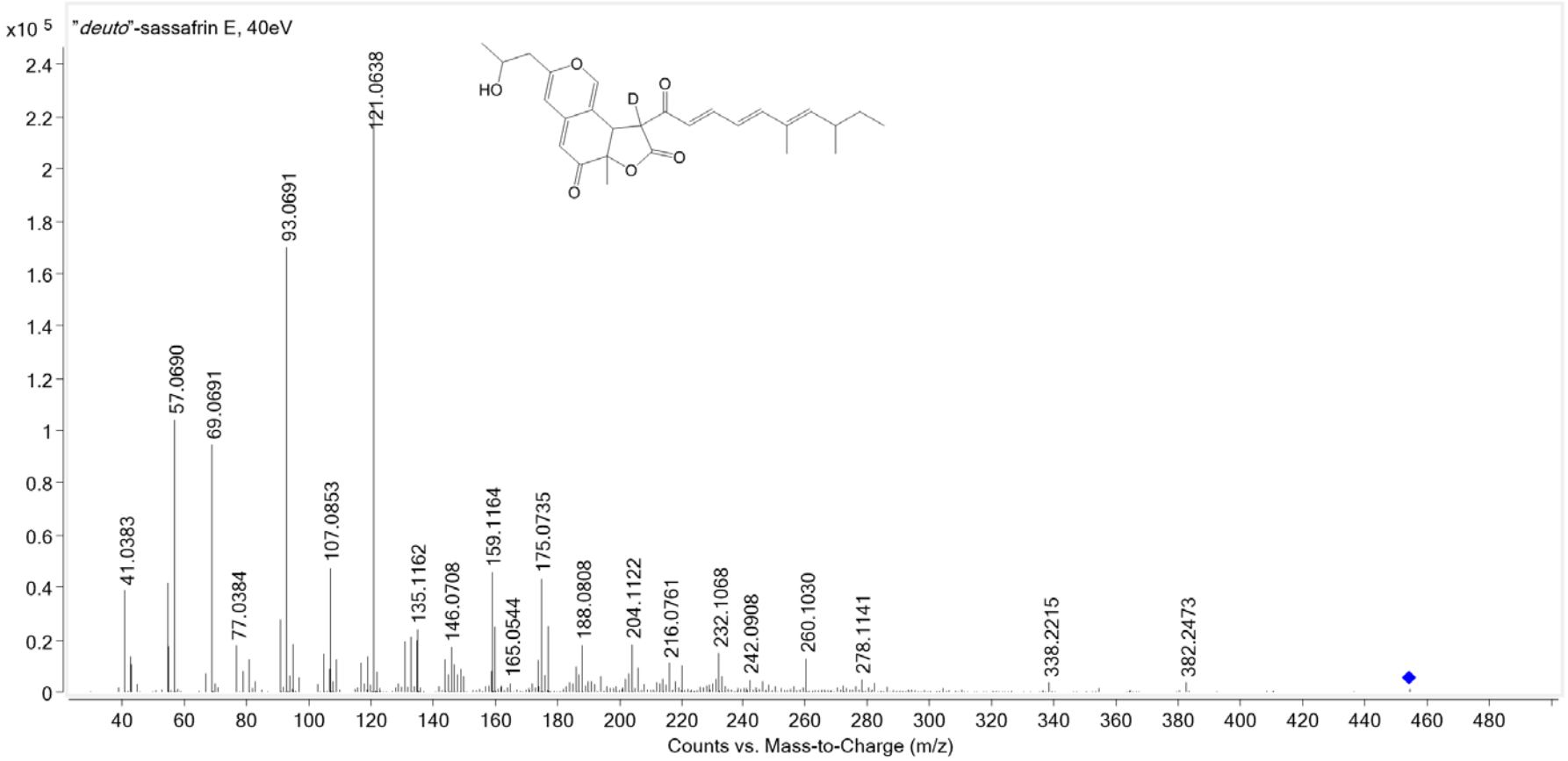


*Figure S3. Tandem mass spectrum for deuterated sassafrin E (**1**) at 40 eV colission energy*

Appendix 5

Structure and genetic origin of novel class of polyketide biomarkers from *Aspergillus brasiliensis*, brasenols

Thomas Isbrandt, Jakob Kræmmer Rendsvig, Rasmus John Norman Frandsen, Jakob Blæsbjerg Hoof*, and Thomas Ostenfeld Larsen*

Department of Biotechnology and Biomedicine, Technical University of Denmark, Kongens Lyngby, Denmark.

*Corresponding authors

Abstract

The black aspergillus, *Aspergillus brasiliensis*, a black aspergillus known to produce compounds such as naphto-γ-pyrones and cyclic peptides. From analysis by ultra-high performance liquid chromatography coupled to diode array detection and high-resolution mass spectrometry (UHPLC-DAD-HRMS), one strain of *A. brasiliensis* was found to be a prolific producer of a novel compound with strong absorption at 350 nm. In this paper, we present the structuras well as the genetic origin of a novel class of compounds, brasenols. The compounds are reduced polyketides linked to 3-hydroxy fatty acids of varying length and saturation via a core lactone. During initial bioactivity tests of one of the analogues, brasenol A1 demonstrated antibacterial properties. Furthermore upon heterologous expression of the gene cluster in *A. nidulans* sporulation was greatly inhibited, likely due to interference of the primary fatty acid metabolism.

Keywords: *Aspergillus brasiliensis*, secondary metabolites, NMR, comparative genomics

Introduction

Fungal secondary metabolites (SM) are a well-known source for interesting compounds; e.g. bioactives for use as pharmaceuticals or pesticides, or pigments for use in the food industry. However, even though many compounds have been discovered based on bioactivity guided approaches, many fungal species have not been thoroughly investigated, thus leaving a great amount of novel compounds still to be found. Additionally, advancements in genome sequencing technologies has made bioinformatics approaches to natural products discovery and pathway elucidation a much more attractive strategy.

Chemical investigation of an extract from the filamentous fungus *A. brasiliensis*, resulted in discovery of a class of polyketide derived compounds not previously described, brasenols. When searching our in-house database, we discovered that these compounds were also produced by other black aspergilli, namely *A. carbonarius* and *A. ellipticus*. This encouraged us to identify the gene cluster responsible for production of the compounds. BLAST analysis of two of the producing fungi, *A. brasiliensis* and *A. carbonaius*, as well as the

non-brasenol producing *A. tubingensis*, was used to identify the responsible polyketide synthase (PKS), allowing identification of the gene cluster. The *brs* gene cluster was determined to encode the PKS, BrsA, a hydrolase, BrsB, an esterase, BrsC, and a transporter, BrsD, and a proposed biosynthetic pathway could be suggested.

Results and Discussion

Isolation and identification of novel metabolite from *A. brasiliensis*

Aspergillus brasiliensis is a producer of known compounds such as aurasperones, pyroranonigrins, aflavinine, tensidols, and malformins[1,2], as well as several compounds that appear to be unique to *A. brasiliensis*. By dereplication of HPLC-DAD-MS data obtained from analysis of plug extracts of *A. brasiliensis* strain type culture (CBS 101740/IBT 21946) cultivated on YES (Yeast Extract Sucrose) growth medium, we detected the presence of a relatively large peak having a mass-to-charge ratio (m/z) of 331.1909 (corresponding to a molecular formula of $C_{20}H_{26}O_4$, DBE = 8, calculated m/z = 331.1904, 1.51 ppm mass accuracy). Moreover, the compound showed a non-familiar UV spectrum with strong absorption at 350 nm, suggesting the presence of a conjugated double bond system (Supplementary Material S1). Adding these observations to the ratio of protons to carbons and of oxygens in the formula, indicated that the compound could be a reduced polyketide, or potentially derived from fatty acids, and we speculated that it was an unknown secondary metabolite.

In order to elucidate the structure of the main compound having the m/z = 331.1904, we cultivated the strain on 200 solid agar plates of YES medium and harvested all biomass and media with ethyl acetate (EtOAc). Using flash chromatography and semi-preparative reversed phase chromatography, we isolated approximately 3 mg of the compound. The pure compound was subjected to one- and two-dimensional NMR spectroscopy and from the ^{13}C spectrum a total of 20 carbon atoms distributed along 19 signals were identified. The ^{1}H spectrum along with the multiplicity edited HSQC could detect all 26 protons, listed in Table 1 (NMR spectra can be found in Supplementary Material S3). DQF-COSY provided correlations for two spin systems; one consisting of atoms **1** through **9**, and one consisting of **1'** through **7'** (Figure 1). From HMBC experiments carbons **8'**, **10**, **11**, and **12** could be linked to the two spin systems, placing them on each side of a central lactone ring. An unusually high chemical shift of 17.6 ppm was determined to originate from an enol, with the proton situated on one of the oxygen atoms of carbons **8'** or **10**. The placement of the enol-proton on the oxygen bonded to **10** rather than **8'** was determined by a W-coupling in the DQF-COSY spectrum from the enol-proton to carbon **9**. It is, however, worth noting that the spectra were recorded in deuterated chloroform, and that in a biological environment, both tautomers are believed to exist in equilibrium. To further confirm the placement of carbons **8'** and **10**, a 1,1-ADEQUATE experiment was recorded, confirming the **7'-8'** and **9-10** linkages. The structure of the compound, which we have named brasenol A1, is shown in Figure 1.

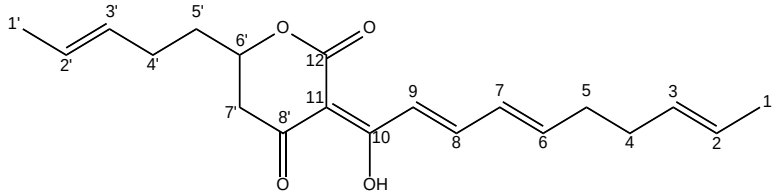


Figure 1. Numbered structure of brasenol A1.

Based on the structural elucidation, we expected the compound to originate from a hexaketide (**1**-**12**) and a fatty acid (**1'-8'**), and in addition to brasenol A1, several putative analogues were identified with similar absorption spectra. These included m/z's 333.2060, 357.2060, 359.2217, 361.2373, 383.2217 and 385.2373, and we speculated these to contain fatty acid moieties of varying lengths and degrees saturation. The literature only contains a few examples of compounds similar to brasenol A1: The fungal compounds alternaric acid[3], fujikurins A-D[4], lachnelluloic acid[5], and CR377[6], and the plant metabolites podoblastin A-C[7] are compounds a with a similar lactone ring structure. Several of these compounds have been reported to be bioactive[3,6,7], Initial bioactivity tests showed brasenol A1 to possess mild antibacterial properties against methicillin-resistant *Staphylococcus aureus* (MRSA) MB5393, with a minimum inhibitory concentration (MIC) of 28.44 µg/mL. However, none of the five mentioned compounds also having the lactone ring structure appeared to display antibacterial activity.

Table 1. ^{1}H and ^{13}C NMR-shifts for brasenol A1.

	Brasenol A1	
#	δ_H	δ_C
1	1.63 (m)	17.6
2	5.45 (m)	126.1
3	5.36 (m)	129.1
4	2.1 (m)	31.3
5	2.25 (q, *J* = 7.2 Hz)	33.1
6	6.27 (m)	147.6
7	6.34 (dd, *J* = 11/14.8 Hz)	129.8
8	7.54 (dd, *J* = 10.8/15 Hz	147.3
9	7.38 (d, *J* = 15 Hz)	122.0
10	-	186.3
10-OH	17.59 (s)	-
11	-	101.7
12	-	164.7
1'	1.63 (m)	17.6

2'	5.45 (m)	126.1
3'	5.36 (m)	129.1
4'	2.15 (m)	27.5
5'	1.64/1.84 (m)	34.2
6'	4.35 (m)	72.9
7'	2.55/2.59 (m)	39.5
8'	-	198.3

77

78 **Identifying backbone biosynthetic machinery for production of brasenols**

79 We hypothesized brasenol A1 to be of polyketide origin, and therefore set out to investigate which genes were
80 responsible for production of the compound. To facilitate the genome mining by using comparative genomics
81 across related species, we examined a handful of the species representing the section Nigri in *Aspergillus* for
82 production of brasenol A1, including *A. niger*, *A. carbonarius*, *A. tubingensis*, and *A. aculeatus*. Interestingly,
83 *A. carbonarius* was also found to be a producer of brasenol A1, as well as the same putative analogues
84 identified in *A. brasiliensis*. In order to identify gene candidates, we performed BlastP analyses using a generic
85 ketide synthase (KS) domain (see Supplementary materials S2) representing the polyketide synthase, against
86 the full set protein sequences of the brasenol producers *A. brasiliensis* and *A. carbonarius* as well as *A.*
87 *tubingensis*, a closely related non-producer.

88 The results from the BLAST analysis allowed us to rank and select candidate proteins according to the display
89 of high homology. The limit was set to more than 50 % identity between PKS candidates in *A. brasiliensis* and
90 *A. carbonarius*, and the candidate should not be encoded by *A. tubingensis*. For this strategy to be valid, we
91 anticipated that the gene cluster was not present in the non-producer, and not just silent. Moreover, the
92 candidate PKS should contain DH, KR and ER domains as well, for full reduction of the growing polyketide
93 chain. Using these criteria as a genome mining approach, only one PKS (ABRA_145702) did not contain any
94 extracurricular activity besides the three reductive domains DH, KR and ER. Other candidates contained either
95 condensation (C) domains, which is characteristic for PKS-nonribosomal peptide synthetase (NRPS) hybrids,
96 or methyl transferase (MT) domains – no methylation was expected in brasenol A1.

97 **Delineating the gene cluster for brasenol production**

98 The PKS candidate was deleted in our *A. brasiliensis akuA*Δ strain and validated by diagnostic PCR (data not
99 shown) following the scheme depicted in Supplementary Material S5. Verified homokaryotic deletion strains
100 were cultivated on YES media at 30 °C for one week followed by plug extraction. The metabolite profile for
101 the mutant was analysed by UHPLC-DAD-HRMS and compared to a reference strain still expressing the gene
102 (Figure 2). In the extracts of the Abra_145702 (named *brsA*) deletion strain, no brasenol was detected, strongly

suggesting the gene to encode a PKS involved in brasenol production. Reintroducing the PKS into the *brsA* deletion strain completely restored production of brasenols, confirming that this gene indeed could be linked to brasenol production. Upon reintroduction of the PKS, the ratio of brasenols was slightly changed, resulting in a mutant producing elevated amount of tentative analogues containing a proposed twelve-carbon fatty acid. We expect this to be a result of the altered regulation of the reintroduced gene, as it was constitutively expressed.

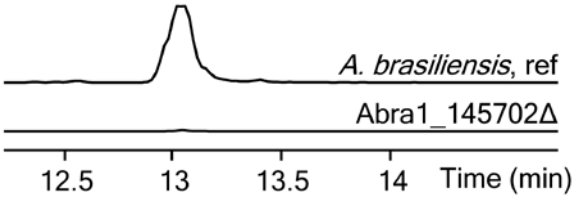


Figure 2. Extracted ion chromatograms of brasenol A1 (m/z=331.1904±0.01) in the A. brasiliensis wildtype and the brsA deletion mutant.

After the successful abolishment of brasenol production, we re-examined the genome sequence in search for nearby genes encoding enzymes predicted to perform aldolase- or esterase-, or hydrolase-like functions in hope of further clarifying the biosynthetic route to brasenols. We reasoned that a synteny analysis for the *brs* locus would give a strong clue to the genes that were involved in the gene cluster. Moreover, if we could find more putative producers of brasenols, we would strengthen this hypothesis. Searching our extracts of Section Nigri species cultivated on solid YES medium revealed that *A. ellipticus* also had the ability to produce brasenol. Subsequently, BlastP analysis for BrsA homologues revealed as expected that the two black Aspergilli, *A. carbonarius* (Protein 209589, 86.3 % Identity) and *A. ellipticus* (Protein 368299, 68.8 % Id.) had candidate PKSs with strong homology to BrsA, as well as *Talaromyces verruculosus* (entry ZTR_03817, 87 % Id.) which also had the genetic material needed to produce a protein highly similar to BrsA. The respective loci sequences for *A. brasiliensis*, *A. carbonarius*, *A. ellipticus* and *T. verruculosus* were retrieved and genes surrounding the PKS annotated. The sequence stretches for both *A. carbonarius* and *A. ellipticus* were relatively short and containing only five and four predicted genes, respectively. However, all four loci were analysed for synteny in EasyFig, and they all shared four core genes (Figure 3).

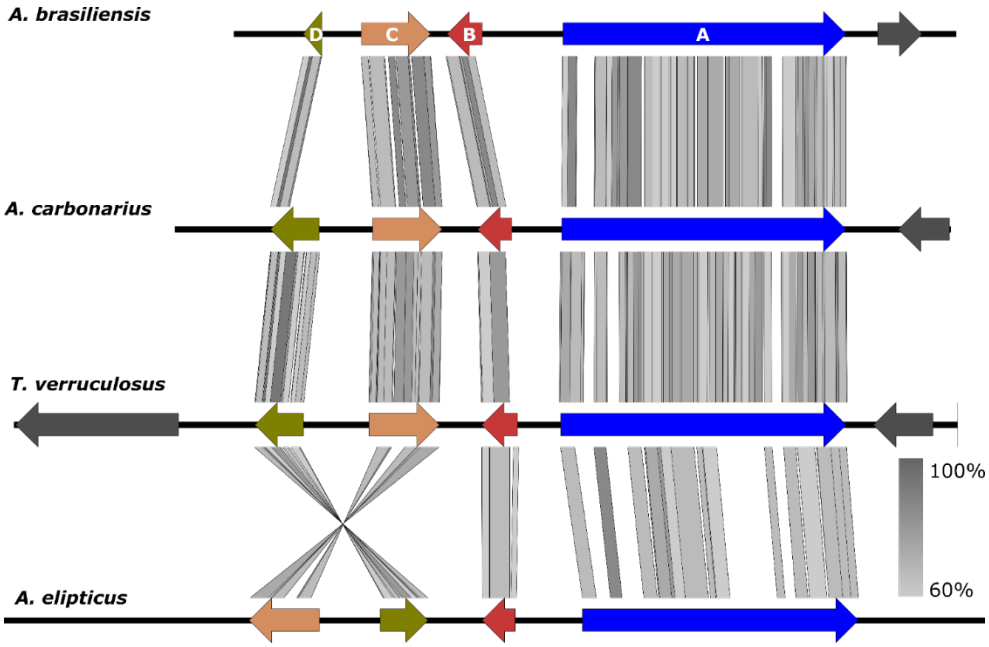


Figure 3. Synteny plot of the loci responsible for brasenol production. Three species showed an identical organization of the gene cluster, and in A. ellipticus the order of brsC and brsD is flipped. Morever, from the plot it is apparent that the gene model of brsD in A. brasiliensis appears to be incorrect and too short.

Deleting these three additional genes individually in *A. brasiliensis* enabled us to study the chemical profiles of these mutants. In the Abra_170790Δ strain, *brsB*Δ, brasenol was completely abolished. The predicted function of the produced protein is an α/β-hydrolase, which we speculate could be responsible for the carbon-carbon bond formation between the PK chain and 3-hydroxy fatty acids. Analysis of the Abra_39116Δ mutant, *brsC*Δ, turned out to also have impaired brasenol production, however with minor quantities of brasenol still produced. This gene product is predicted to be an acyl esterase, and we speculate that it could be responsible for the ring formation of the lactone. The last gene deletion of the four candidates Abra_39117, *brsD*, an MFS transporter, showed colonies that appeared much smaller when compared to the other three candidates. The resulting deletion mutant was affected in colony diameter and delayed in sporulation on MM and production of brasenol was still detected, but reduced (data not shown).

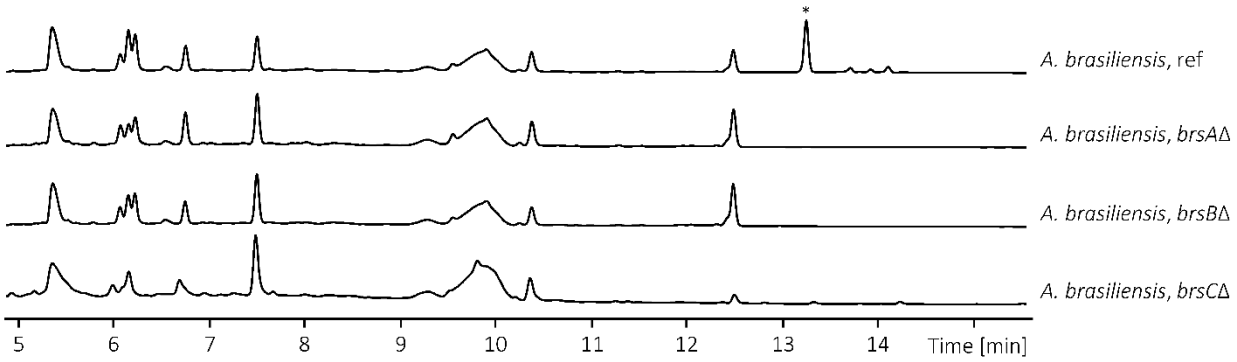


Figure 4. Deletion of brsA, brsB and brsC in A. brasiliensis resulted in either termination or heavy downregulation of brasenol production.

Heterologous production of brasenol

We decided to examine the minimum genetic requirement for brasenol production by heterologous expression of the two genes, *brsA* and *brsB*, which showed to be necessary for production in *A. brasiliensis*. The expression of either *brsA* or *brsB* did not result in any changes in the resulting *A. nidulans* mutant strains compared to the reference strain, as neither brasenol or any potential precursors (aliphatic hexaketide or fatty acid) were identified upon chemical analysis.

However, co-expression of *brsA* and *brsB* under the control of the H3/H4 dual promoter from the same locus now resulted in dramatic changes to *A. nidulans*. Firstly, massive production of brasenol A1 was observed as shown in Figure 5, indicating that the hydrolase, BrsB, is necessary for production of brasenols. Secondly, the mutant displayed a significantly altered morphology, strikingly with no production of spores and hampered growth. In addition to brasenol A1, the same analogues as observed in the *A. brasiliensis* wildtype were also detected in the *A. nidulans* overexpression mutant. Furthermore when BrnA and BrnB were heterologously expressed in both *A. sydowii* and *A. oryzae*, both experiments resulted in similar production of brasenols (data not shown). In addition to the high production levels of brasenols in *A. nidulans*, it also had an impact on some of the prominent endogenous compounds, most significantly austinol, dehydroaustinol, and sterigmatocystin, which based on LC-MS analysis were determined to be downregulated. Similarly, arugosin A was upregulated, along with a few unknown compounds which were not further investigated.

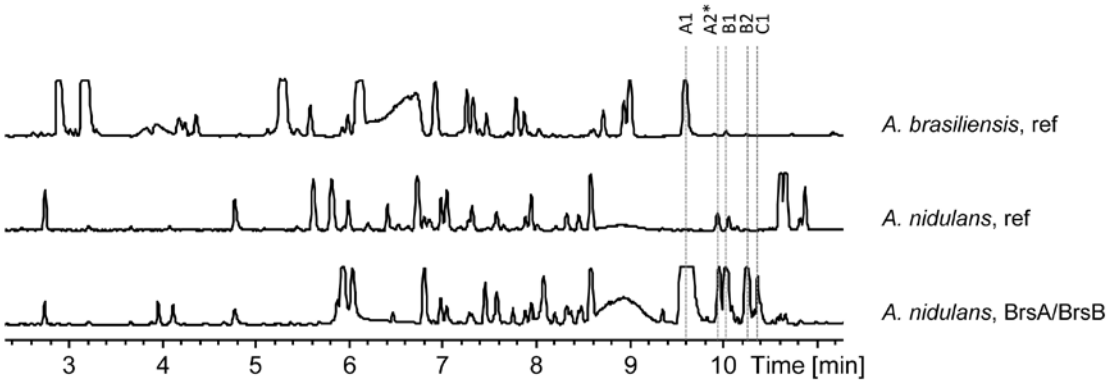


*Figure 5. Base peak chromatograms showing production of brasenols in A. brasilisensis and in the A. nidulans BrsA/BrsB overexpression mutant. * indicates a putative analogue (m/z=333.2060) not purified.*

In order to verify the structure of the proposed brasenol analogues, the *A. nidulans* mutant expressing BrsA and BrsB was inoculated on minimal medium for seven days. From 30 plates, 26.7 mg of brasenol A1 was obtained, as well as 1.2 mg of mz357 and 0.5 mg of a 1:1 mixture of mz359 and mz383. NMR spectra were recorded and the structures of the four compounds were elucidated and are shown in Figure 6. Chemical shifts of analogues B1, B2 and C1 can be found in Supporting Information S4 along with the recorded DQF-COSY and HSQC spectra (Supporting Information S3).

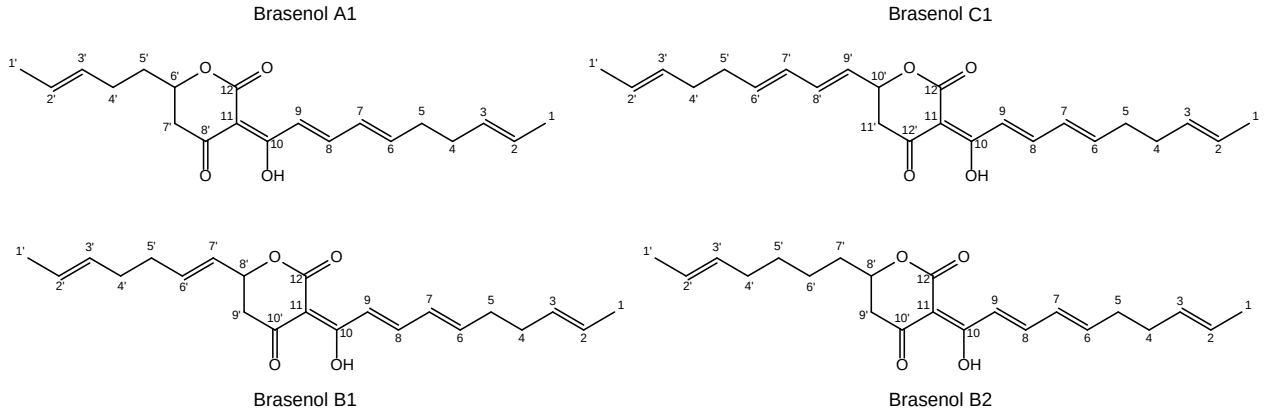


Figure 6. Structures of the four brasenol species characterised in this study.

Following successful expression of the brasenol gene cluster in several different hosts, and the isolation of four analogues, we were able to suggest a biosynthetic pathway for the brasenols, as shown in Figure 7. The compounds are made of two distinct carbon chains; a constant, twelve carbon reduced polyketide, and a 3-hydroxy fatty acid of varying chain length. This hypothesis was supported by the characterisation of the three analogues containing respectively two and four carbons more than brasenol A1. The PKS BrsA is responsible for biosynthesis of the constant twelve-carbon PK chain present in all the isolated brasenol species. The hydrolase BrsB acts by fusing the PK to various types of 3-hydroxy fatty acids via Claisen-like condensations with mainly C8, but also C10 and C12 fatty acids. Release and simultaneous ring formation of the lactone is

proposed to be catalysed by the esterase BrsC. The esterase has homologues in *A. nidulans* (54 % aa ID, 67,5 % positives), *A. sydowii* and *A oryzae*, and we speculate that this is why BrsC is not required for heterologous production of brasenols in these species.

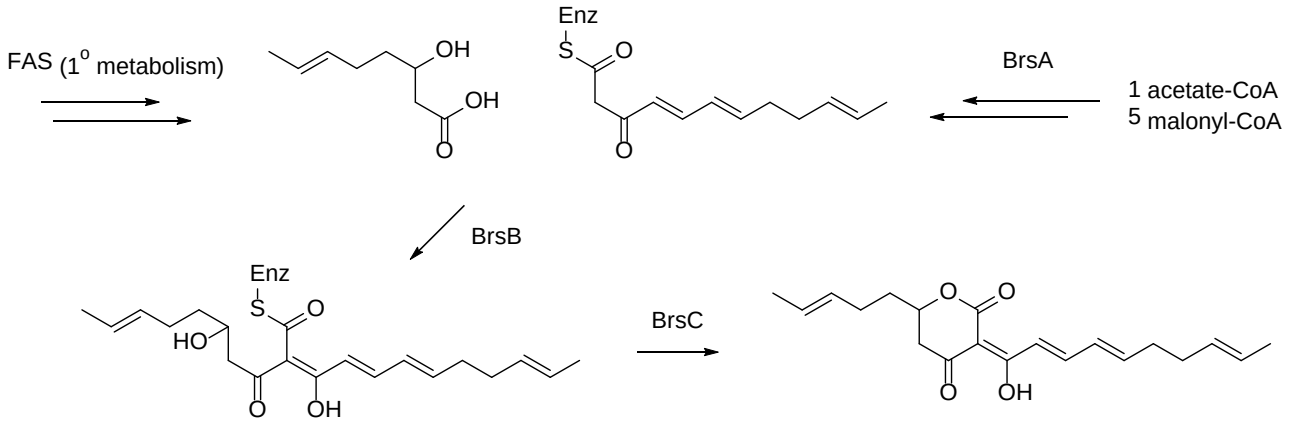


Figure 7. Proposed biosynthetic pathway of brasenols, exemplified with brasenol A1.

It is well known that fatty acids are extremely important in metabolic regulation in nearly all organisms, including fungi. Oxylipins are a type of lipids that play an essential role in various stages of regulation, development and communication, and by interfering with these signalling molecules, adverse effects have been reported to occur.[8–10] Oxylipins are quite diverse with regard to size and structure, with one group, the 3-hydroxy oxylipins, seemingly being involved during ascospore production.[11] Furthermore, 3-hydroxy derivatives of both octanoic acid and decanoic acid were found to be involved in flocculation in *Saccharomyces cerevisiae*, possibly by assisting the adhesion between the cells.[12] The observed effects on sporulation in this study, in conjunction with reported research on fatty acid involvement in fungal reproduction shows a clear link between the two. The exact mechanism by which the brasenol genes are affecting the heterologous hosts can so far only be speculated upon. However, the excessive production of brasenols, could indicate that the foreign gene leech on products usually used for other parts of metabolism, thereby disturbing the ability to grow and sporulate. Future work is needed to get a better understanding of the exact interactions causing these effects, since this and several other studies have shown a clear correlation between the fatty acid metabolism and a whole range of metabolic functions in filamentous fungi as well as other organisms.

Experimental section

Chemical analysis. For metabolite profile analysis of *A. brasiliensis* (IBT21946), the fungus was inoculated by three point stab, and grown on YES medium for 5 days. Plug extractions were performed by taking 5 plugs of 6 mm diameter across a colony. The plugs were transferred to Eppendorf tubes and extracted with 800µL of a 3:1 mixture of ethyl acetate and iso-propanol, with 1% (v/v) formic acid (FA), for one hour with sonication. Following sonication, the extraction liquid was decanted to new Eppendorf tubes, and the solvent was

evaporated under a gentle stream of nitrogen gas at 30°C. The dried extracts were re-dissolved in 400µL methanol (MeOH) with sonication, and centrifuged for 3 min at 13500 rpm to avoid any spores or other particles in the sample. For *A. nidulans* plug extractions the same procedure was used with the exception of the growth medium which was minimal medium.

Samples, including fractions from purification, were analysed by UHPLC-DAD-MS on an Agilent Infinity 1290 UHPLC system (Agilent Technologies, Santa Clara, CA, USA) equipped with a diode array detector. Separation was obtained on an Agilent Poroshell 120 phenyl-hexyl column (2.1 × 250 mm, 2.7 µm) with a linear gradient consisting of water (A) and acetonitrile (B) both buffered with 20 mM formic acid, starting at 10% B and increased to 100% in 15 min where it was held for 2 min, returned to 10% in 0.1 min and remaining for 3 min (0.35 mL/min, 60 °C). An injection volume of 1 µL was used. MS detection was performed in positive detection mode on an Agilent 6545 QTOF MS equipped with Agilent Dual Jet Stream electrospray ion source with a drying gas temperature of 250 °C, gas flow of 8 L/min, sheath gas temperature of 300 °C and flow of 12 L/min. Capillary voltage was set to 4000 V and nozzle voltage to 500 V. Mass spectra were recorded at 10, 20 and 40 eV as centroid data for *m*/*z* 85–1700 in MS mode and *m*/*z* 30–1700 in MS/MS mode, with an acquisition rate of 10 spectra/s. Lock mass solution in 70:30 methanol:water was infused in the second sprayer using an extra LC pump at a flow of 15 µL/min using a 1:100 splitter. The solution contained 1 µM tributylamine (Sigma-Aldrich) and 10 µM Hexakis(2,2,3,3-tetrafluoropropoxy)phosphazene (Apollo Scientific Ltd., Cheshire, UK) as lock masses. The $[M + H]^+$ ions (*m*/*z* 186.2216 and 922.0098 respectively) of both compounds was used.

Purification of compounds. Purification of brasenol A1 was done by cultivating *A. brasiliensis* (IBT21946) on 200 plates of YES medium for 7 days. The mycelia and agar was extracted two times with 1% FA in EtOAc, first for one hour with sonication, followed by extraction overnight.

The crude extract was subjected to degreasing and –sugaring, by dissolving in a minimum of 90% MeOH:Milli-Q H_2O and extracting with an equal amount of n-heptane. MQ H_2O was added to the MeOH phase to give a 1:1 mixture of MeOH and water and the phase was extracted with dichloromethane (DCM). Each of the three phases (aqueous, n-heptane, and DCM) was analysed by LC-MS.

The brasenol containing DCM phase was fractioned by stepwise reversed phase (RP) chromatography (C18) on an Isolera One flash chromatography system from Biotage. MeOH and MQ H_2O was used as mobile phase and the elution was achieved in steps 10% MeOH (10-100% MeOH), followed by a 100% acetone fraction containing the compound of interest.

For final purification, a semi-preparative automated Gilson 322 pump module with a Gilson 172 DAD was used. Separation was achieved on a Phenomenex Luna Gemini C6-Ph column using an acetonitrile:H_2O

gradient starting at 70% going to 87% in 17 minutes. Fractions were automatically collected based on absorption at 350 nm.

Brasenol analogues (B1, B2 and C1) were purified from the BrnA/BrnB overexpressing *A. nidulans* strain grown on 30 plates of YES media. The plates containing both media and biomass were extracted with EtOAc and the extract was subjected to semi-preparative HPLC on a Waters 600 Controller with a 996 PDA detector, using a Luna II C18 column (250 x 10 mm, 5 µm, Phenomenex) and a gradient of acetonitrile and water, acidified with 50 ppm TFA as mobile phase.

NMR. NMR spectra were recorded on a Bruker Avance 800 MHz spectrometer located at the Danish Instrument Centre for NMR Spectroscopy of Biological Macromolecules at Carlsberg Laboratory (Now located at the Department of Chemistry at The Technical University of Denmark). Spectra were acquired using standard pulse sequences. The deuterated solvent was $CDCl_3$ and signals were referenced by solvent signals for $CDCl_3$ at $\delta H = 7.26$ ppm and $\delta C = 77.16$ ppm. The NMR data was processed in MestReNova V.10.0.2–15465. Chemical shifts are reported in ppm (δ) and scalar couplings are reported in hertz (Hz). The sizes of the J coupling constants in the tables are the experimentally measured values from the 1D 1H and DQF-COSY spectra. There are minor variations in the measurements, which may be explained by the uncertainty of J and the spectral digital resolution.

Strains and media. An *akuA*Δ deletion strain was used for all gene-targeting experiments in *A. brasiliensis* [Theobald et al., Submitted for publication], and this strain is based on the type culture (WT) *A. brasiliensis* (CBS 101740/IBT 21946)[1]. Genomic DNA (gDNA) from WT *A. brasiliensis* and *A. nidulans* IBT 29539[13] was isolated via FastDNA SPIN Kit for Soil DNA extraction kit (MP Biomedicals, USA). Sequence information was obtained from v1.0 *A. brasiliensis* assembly from Joint Genome Institute, JGI[14]. BlastP analysis was performed at JGI genome portal as well as at NCBI and the Aspergillus Genome Database (AspGD).

All *A. brasiliensis* strains were cultivated at 30ºC on minimal medium (MM), supplemented with 10 mM uridine if required for growth. The minimal medium (MM), transformation media (TM) and media for *pyrG* counter-selection (MM+5-FOA) were as described by Nødvig et al., 2015[15]. Yeast extract sucrose (YES, Frisvad and Samson 2004) growth media was used for chemical analysis. Chemical competent *Escherichia coli* DH5α were applied for vector assembly and plasmid propagation at 37ºC, and *E. coli* cultivations were carried out in Lumia Broth (LB) media (1% Bacto tryptone, 0.5% Bacto yeast extract, 1% NaCl, pH 7.0) supplemented with 0.1% ampicillin. All solid media were supplied with 2% agar.

Construction of mutant strains. PCR conditions for cloning-fragment amplification and USER-cloning procedure were as described in Hansen et al., 2011[16]. All USER cassettes were based on PacI/Nt.BbvCI sites. All primers were synthesized by Integrated DNA technologies and listed in Supporting Information S6, Table

S2. DNA fragments for USER cloning and plasmids were purified using illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare Life Sciences) and GenElute Plasmid Miniprep Kit (Sigma-Aldrich), respectively, according to manufacturer's instructions. For deleting all genes in the *brs* locus, see Figure 3, up- and downstream sequences flanking the coding sequences of the genes were amplified by PCR (Figure S2) and USER cloned into the gene-targeting vector pU2002 (pAC3) that harbours *pyrG* from *A. fumigatus* flanked by a direct repeat sequence (vector backbone described by Hansen et al., 2011). Gene-targeting plasmids were linearized and verified by enzymatic digestion with SwaI, according to manufacturer's instructions (NEB). All gene-targeting construct were added to protoplasts of the *akuA*Δ *pyrG1* strain selecting for only *pyrG* as described by Nødvig et al., 2015[15]. Homokaryosis and deletion of geneswas verified by diagnostic tissue-PCR as described in Nødvig et al, 2015[15] and Figure S5. Protoplastation and transformation of the *akuA*Δ *pyrG1* strain (BRA10) were and conducted as described in Nielsen et al., 2006[17] and Nødvig et al., 2015[15], respectively.

References

(1) Varga, J.; Kocsube, S.; Toth, B.; Frisvad, J. C.; Perrone, G.; Susca, A.; Meijer, M.; Samson, R. A. *Int. J. Syst. Evol. Microbiol.* **2007**, *57* (8), 1925.

(2) Lamboni, Y.; Nielsen, K. F.; Linnemann, A. R.; Gezgin, Y. K.; Hell, K.; Nout, M. J. R.; Smid, E. J.; Tamo, M.; Van Boekel, M. A. J. S.; Hoof, J. B.; Frisvad, J. C. *PLoS One* **2016**, *11* (10), 1.

(3) Brian, P. W.; Curtis, P. J.; Hemming, H. G.; Unwin, C. H.; Wright, J. M. *Nature* **1949**, *164* (4169), 534.

(4) von Bargen, K. W.; Niehaus, E.-M.; Krug, I.; Bergander, K.; Würthwein, E.-U.; Tudzynski, B.; Humpf, H.-U. *J. Nat. Prod.* **2015**, *78* (8), 1809.

(5) Ayer, W. A.; Villar, J. D. F. *Can. J. Chem.* **1985**, *63* (6), 1161.

(6) Brady, S. F.; Clardy, J. *J. Nat. Prod.* **2000**, *63* (10), 1447.

(7) Miyakado, M.; Inoue, S.; Tanabe, Y.; Watanabe, K.; Ohno, N.; Yoshioka, H.; Mabry, T. J. *Chem. Lett.* **1982**, *11* (10), 1539.

(8) Tsitsigiannis, D. I.; Keller, N. P. *Trends Microbiol.* **2007**, *15* (3), 109.

(9) Brodhun, F.; Feussner, I. *FEBS J.* **2011**, *278* (7), 1047.

(10) Kock, J. L. F.; Strauss, C. J.; Pohl, C. H.; Nigam, S. *Prostaglandins Other Lipid Mediat.* **2003**, *71* (3–4), 85.

(11) Gessler, N. N.; Filippovich, S. Y.; Bachurina, G. P.; Kharchenko, E. A.; Groza, N. V.; Belozerskaya, T. A. *Appl. Biochem. Microbiol.* **2017**, *53* (6), 628.

(12) Kock, J. L. F.; Venter, P.; Smith, D. P.; Van Wyk, P. W. J.; Botes, P. J.; Coetzee, D. J.; Pohl, C. H.; Botha, A.; Riedel, K. H.; Nigam, S. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* **2000**, *77* (4), 401.

(13) Nielsen, J. B.; Nielsen, M. L.; Mortensen, U. H. *Fungal Genet. Biol.* **2008**, *45* (3), 165.

(14) de Vries, R. P.; Riley, R.; Wiebenga, A.; Aguilar-Osorio, G.; Amillis, S.; Uchima, C. A.; Anderluh, G.; Asadollahi, M.; Askin, M.; Barry, K.; Battaglia, E.; Bayram, Ö.; Benocci, T.; Braus-Stromeyer, S. A.; Caldana, C.; Cánovas, D.; Cerqueira, G. C.; Chen, F.; Chen, W.; Choi, C.; Clum, A.; dos Santos, R. A. C.; Damásio, A. R. de L.; Diallinas, G.; Emri, T.; Fekete, E.; Flipphi, M.; Freyberg, S.; Gallo, A.; Gournas, C.; Habgood, R.; Hainaut, M.; Harispe, M. L.; Henrissat, B.; Hildén, K. S.; Hope, R.; Hossain, A.; Karabika, E.; Karaffa, L.; Karányi, Z.; Kraševec, N.; Kuo, A.; Kusch, H.; LaButti, K.; Lagendijk, E. L.; Lapidus, A.; Levasseur, A.; Lindquist, E.; Lipzen, A.; Logrieco, A. F.; MacCabe, A.; Mäkelä, M. R.; Malavazi, I.; Melin, P.; Meyer, V.; Mielnichuk, N.; Miskei, M.; Molnár, Á. P.; Mulé, G.; Ngan, C. Y.; Orejas, M.; Orosz, E.; Ouedraogo, J. P.; Overkamp, K. M.; Park, H.-S.; Perrone, G.; Piumi, F.; Punt, P. J.; Ram, A. F. J.; Ramón, A.; Rauscher, S.; Record, E.; Riaño-Pachón, D. M.; Robert, V.; Röhrig, J.; Ruller, R.; Salamov, A.; Salih, N. S.; Samson, R. A.; Sándor, E.; Sanguinetti, M.; Schütze, T.; Sepčić, K.; Shelest, E.; Sherlock, G.; Sophianopoulou, V.; Squina, F. M.; Sun, H.; Susca, A.; Todd, R. B.; Tsang, A.; Unkles, S. E.; van de Wiele, N.; van Rossen-Uffink, D.; Oliveira, J. V. de C.; Vesth, T. C.; Visser, J.; Yu, J.-H.; Zhou, M.; Andersen, M. R.; Archer, D. B.; Baker, S. E.; Benoit, I.; Brakhage, A. A.; Braus, G. H.; Fischer, R.; Frisvad, J. C.; Goldman, G. H.; Houbraken, J.; Oakley, B.; Pócsi, I.; Scazzocchio, C.; Seiboth, B.; VanKuyk, P. A.; Wortman, J.; Dyer, P. S.; Grigoriev, I. V. *Genome Biol.* **2017**, *18* (1), 28.

(15) Nødvig, C. S.; Nielsen, J. B.; Kogle, M. E.; Mortensen, U. H. *PLoS One* **2015**, *10* (7), e0133085.

(16) Hansen, B. G.; Salomonsen, B.; Nielsen, M. T.; Nielsen, J. J. B. J. B.; Hansen, N. B.; Nielsen, K. F.; Regueira, T. B.; Nielsen, J. J. B. J. B.; Patil, K. R.; Mortensen, U. H. *Appl. Environ. Microbiol.* **2011**, *77* (9), 3044.

(17) Nielsen, M. L.; Albertsen, L.; Lettier, G.; Nielsen, J. B.; Mortensen, U. H. *Fungal Genet. Biol.* **2006**, *43* (1), 54.

Supplementary Material for

Structure and genetic origin of novel class of polyketide biomarkers from *Aspergillus brasiliensis*, brasenols

Thomas Isbrandt, Jakob Blæsbjerg Hoof, Thomas Ostenfeld Larsen

Department of Biotechnology and Biomedicine, Technical University of Denmark, Kongens Lyngby, Denmark

Contents:

S1. Brasenol A1 absorbtion spectrum:

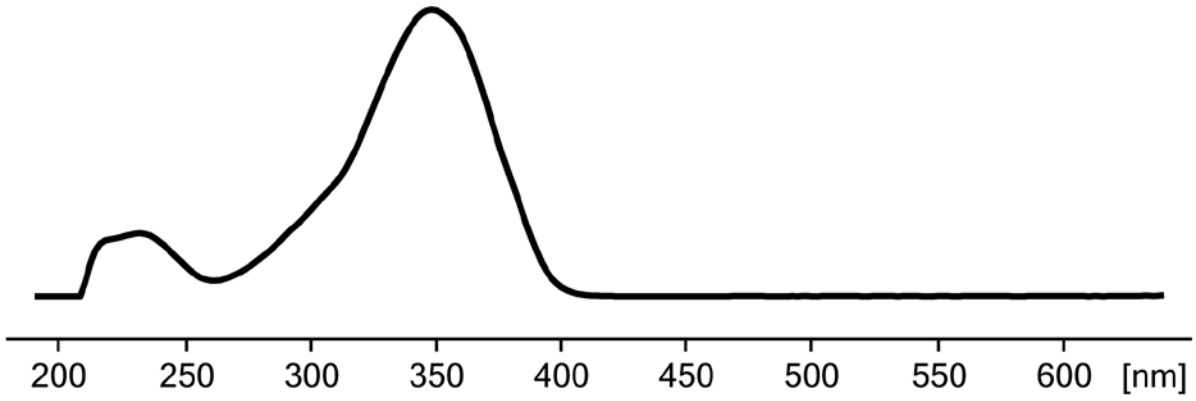


Figure S1. UV spectrum of brasenol A1

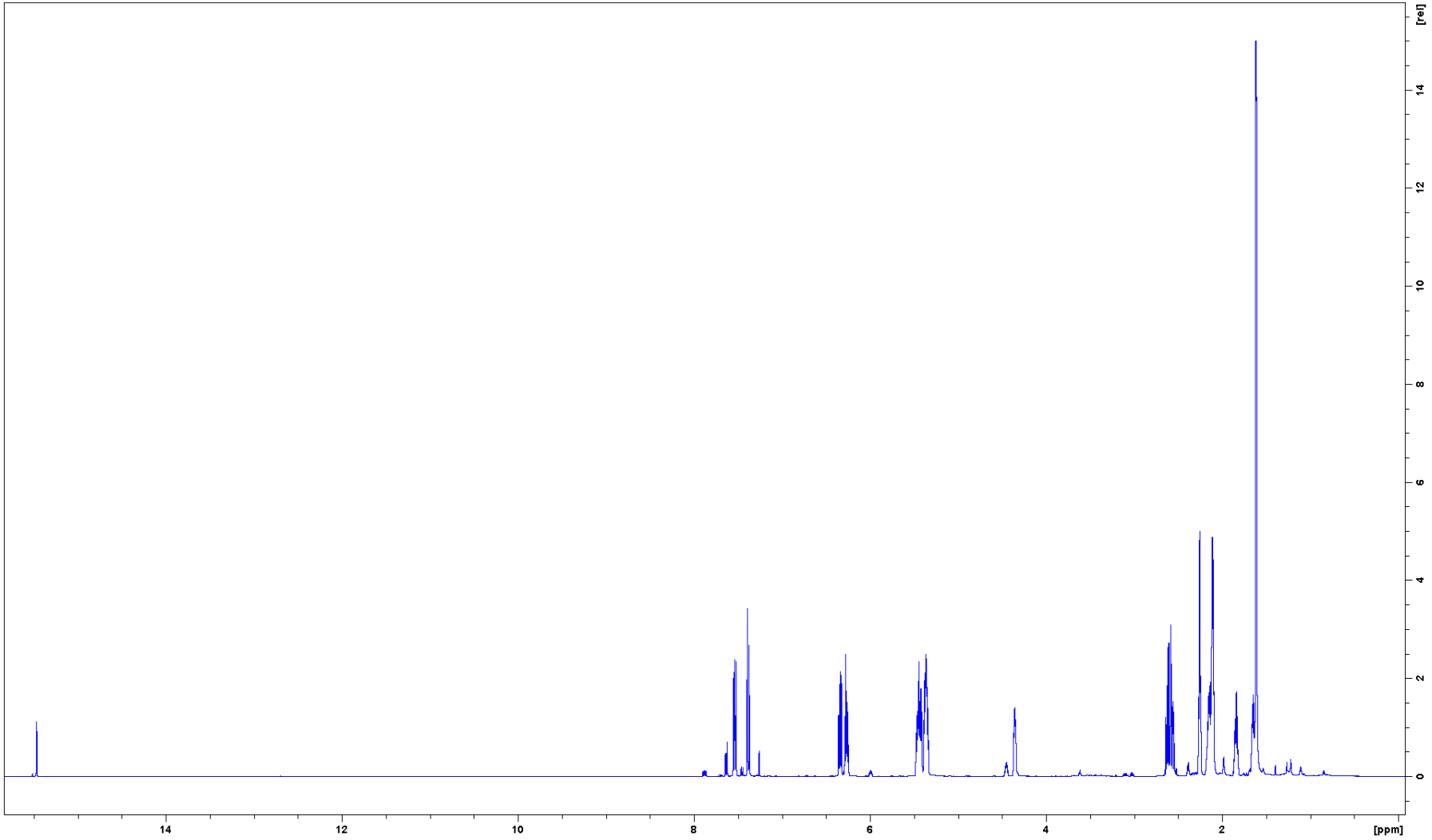
S2. KS sequence used for identification of BrsA gene candidates:

IAIIGSACRFPGDSSSPSKLWDLLKAPRDLLTEVPSNRYNADAFYHADSKHHGTTNVRHSYFLSEDPASFDNNFFNIQPGEAE
AIDPQQRLLMEVVYQGLCASGQTIEGLRGSSTSVYVGVCDDWNGILTRDLEVFPQYGATGMARSIMSNRISYFFDWHGPS
MTIDTACSSSLVAVHQAIQTLRSGESQVAIAAGANLILTPGMYVAESKLSMLSPSGRSKMWDQDVNGYARGEGIAAVVLK
PLSAAIRDNDHIDCIIRATGVNQDGRTPGLTMPSAAAQADLIRSTYARAGLDINKPEDRPQFFHAHGTGTPAGDPRTKA

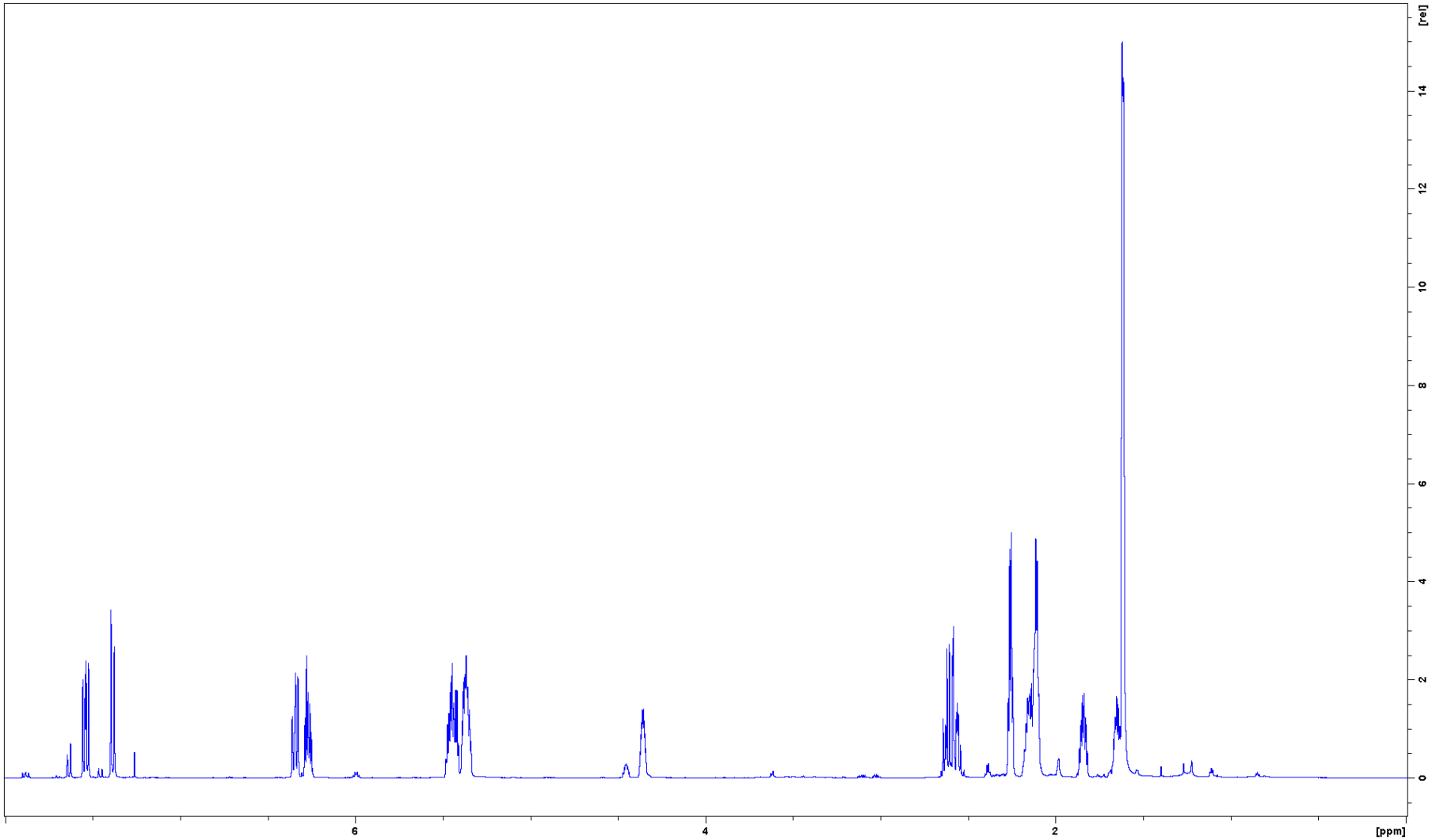
S3. NMR spectra for all brasenols. All spectra were recorded in CD_3Cl at 800 MHz for ^{1}H and 200 MHz for ^{13}C:

Brasenol A1

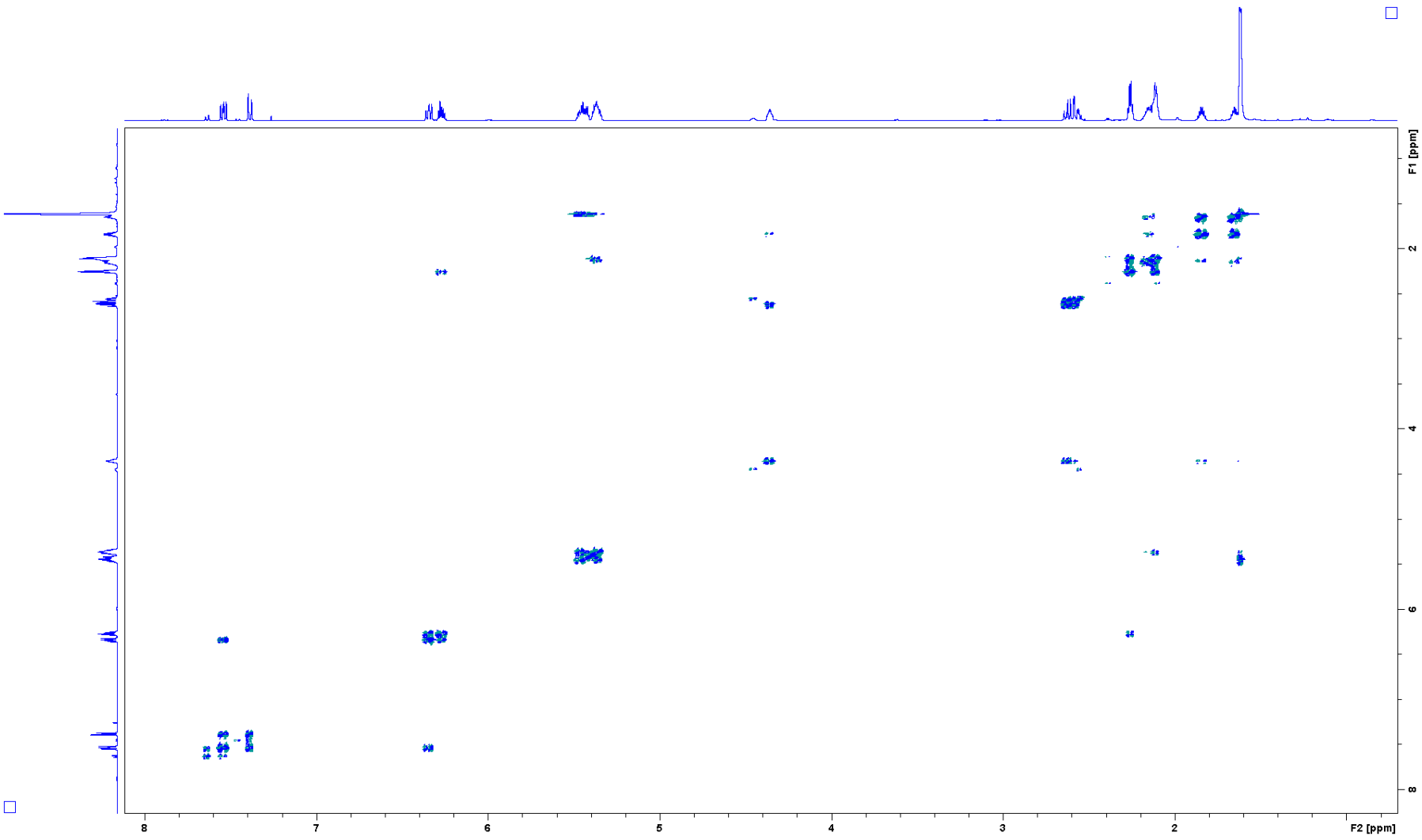
^{1}H-NMR (full range)



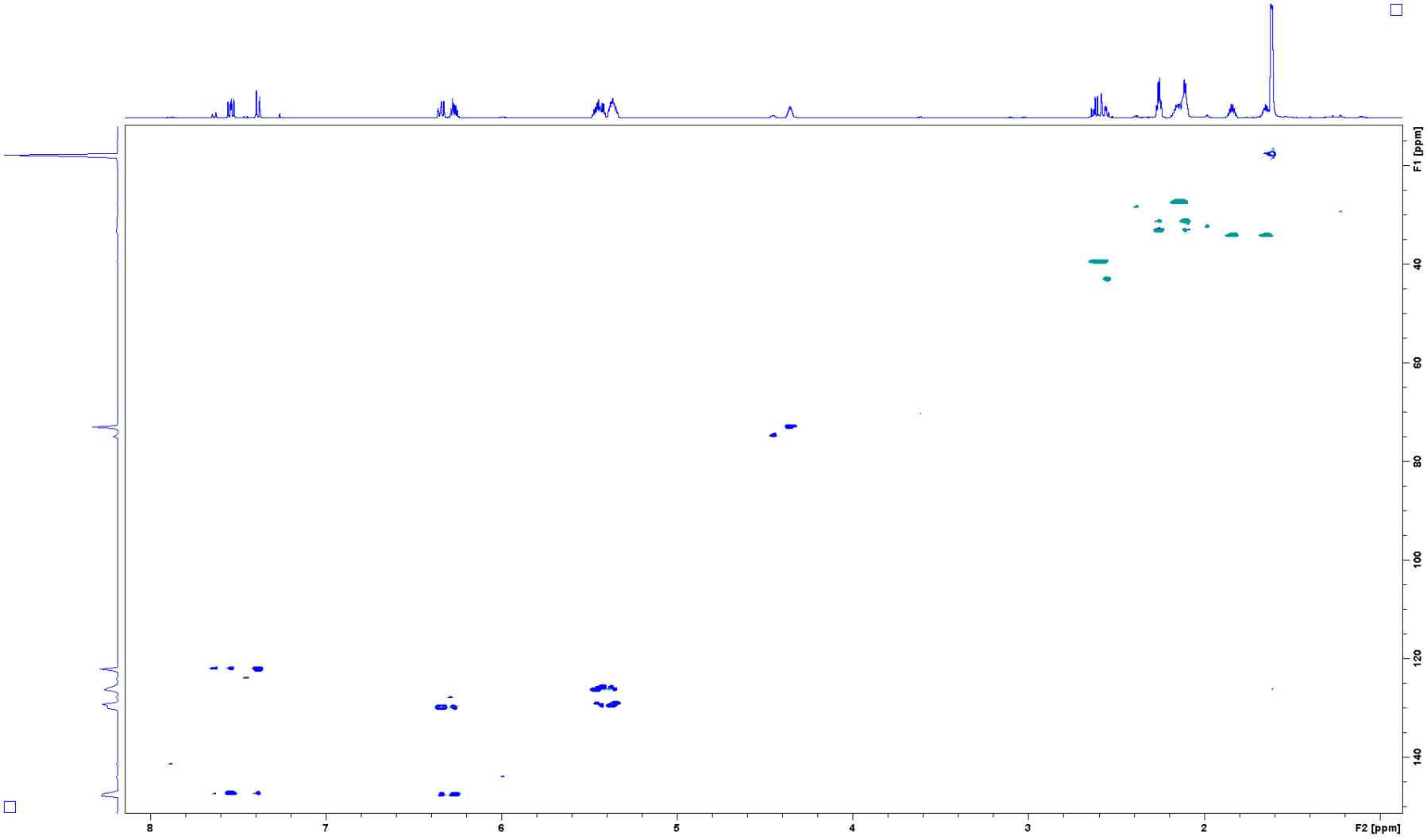
^{1}H-NMR (0-8 ppm)



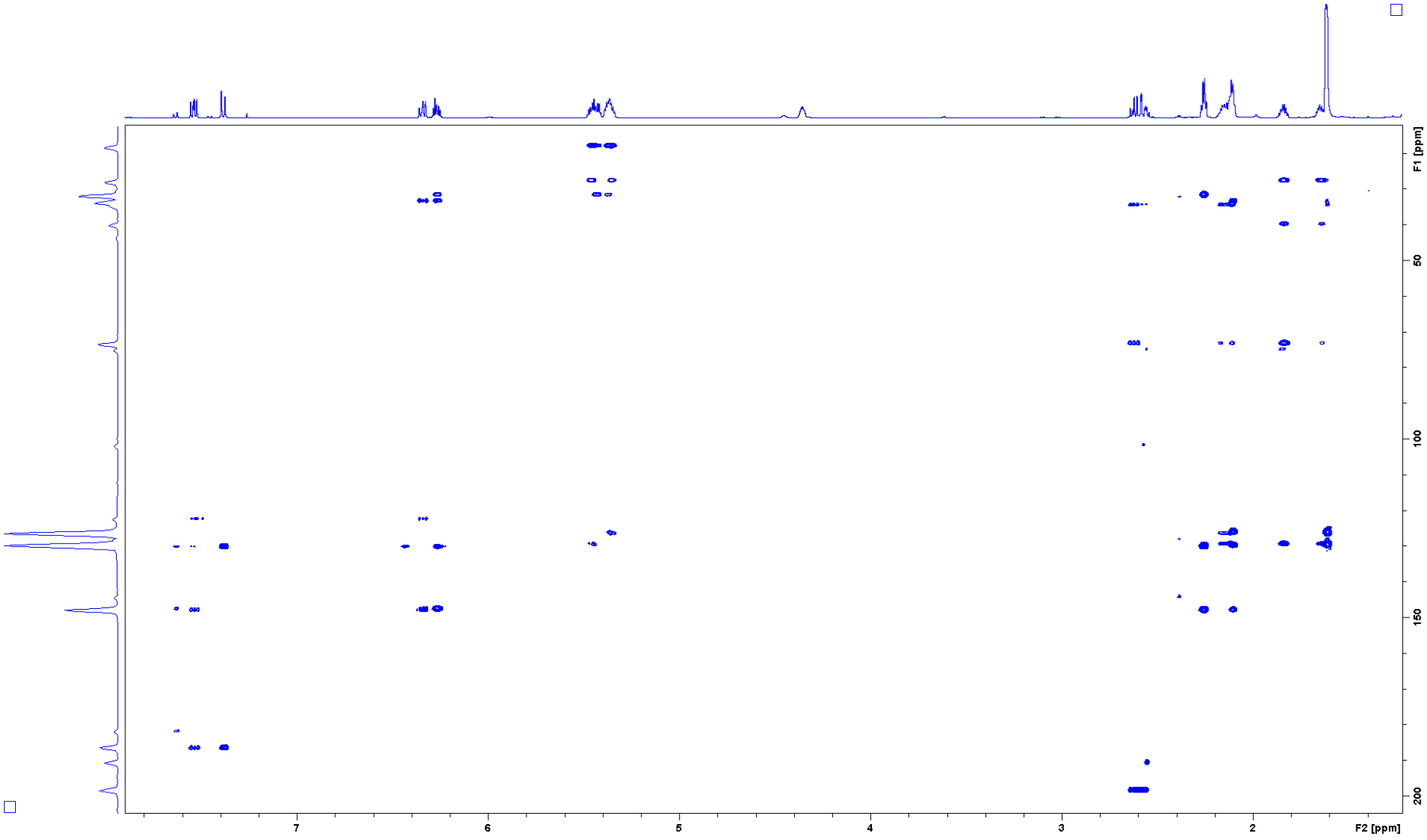
DQF-COSY



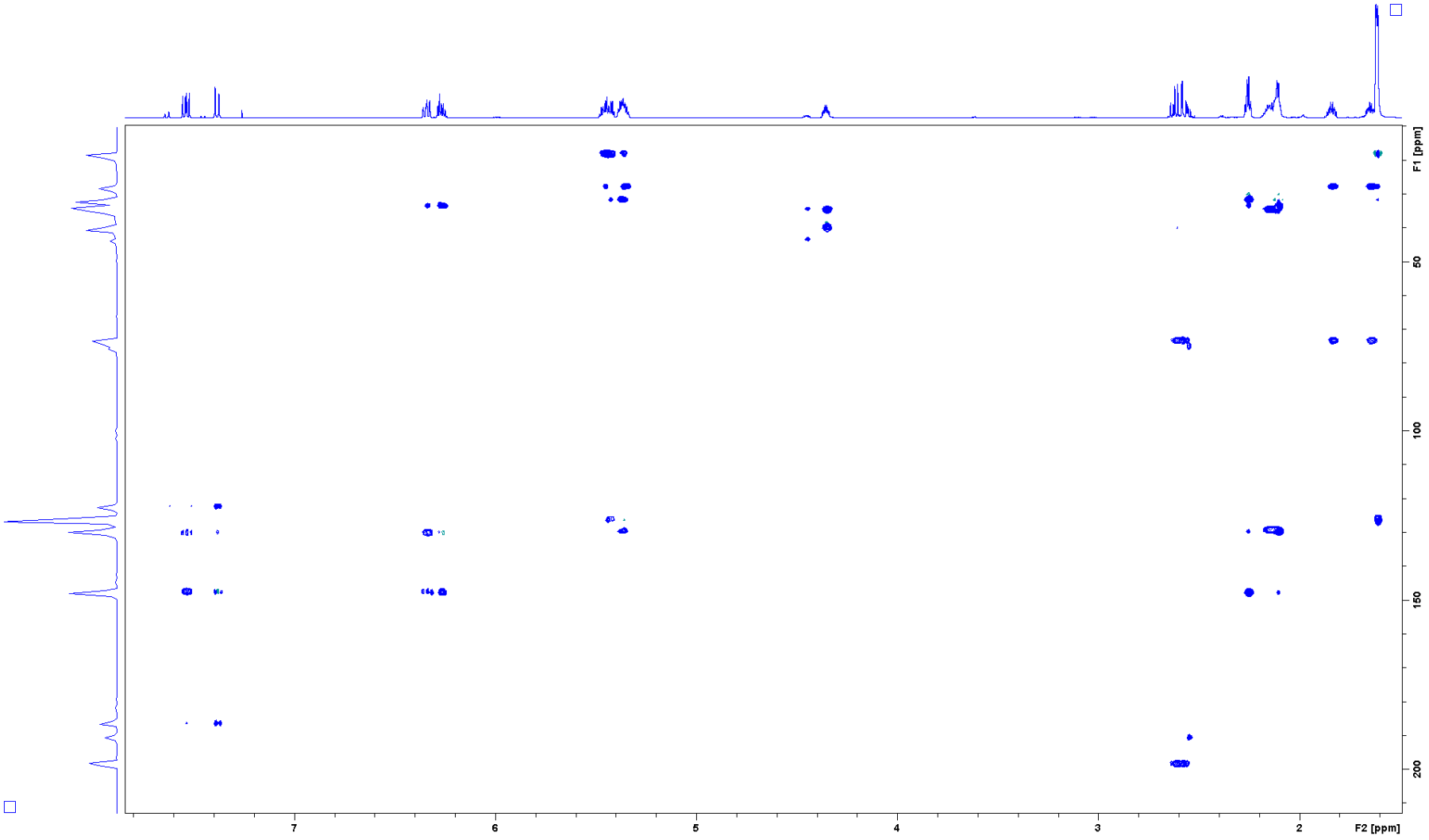
edHSQC



HMBC

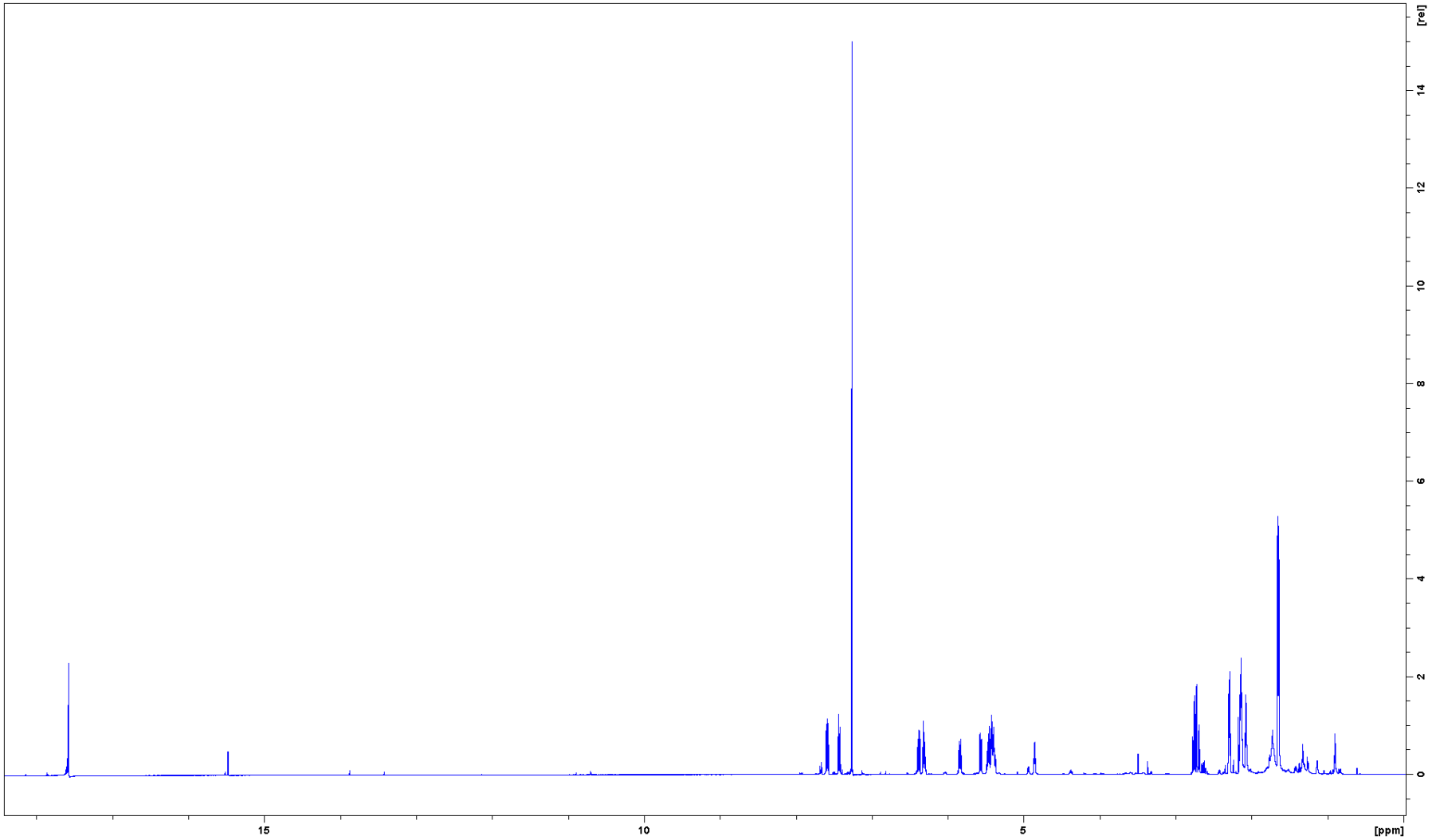


1,1-ADEQUATE

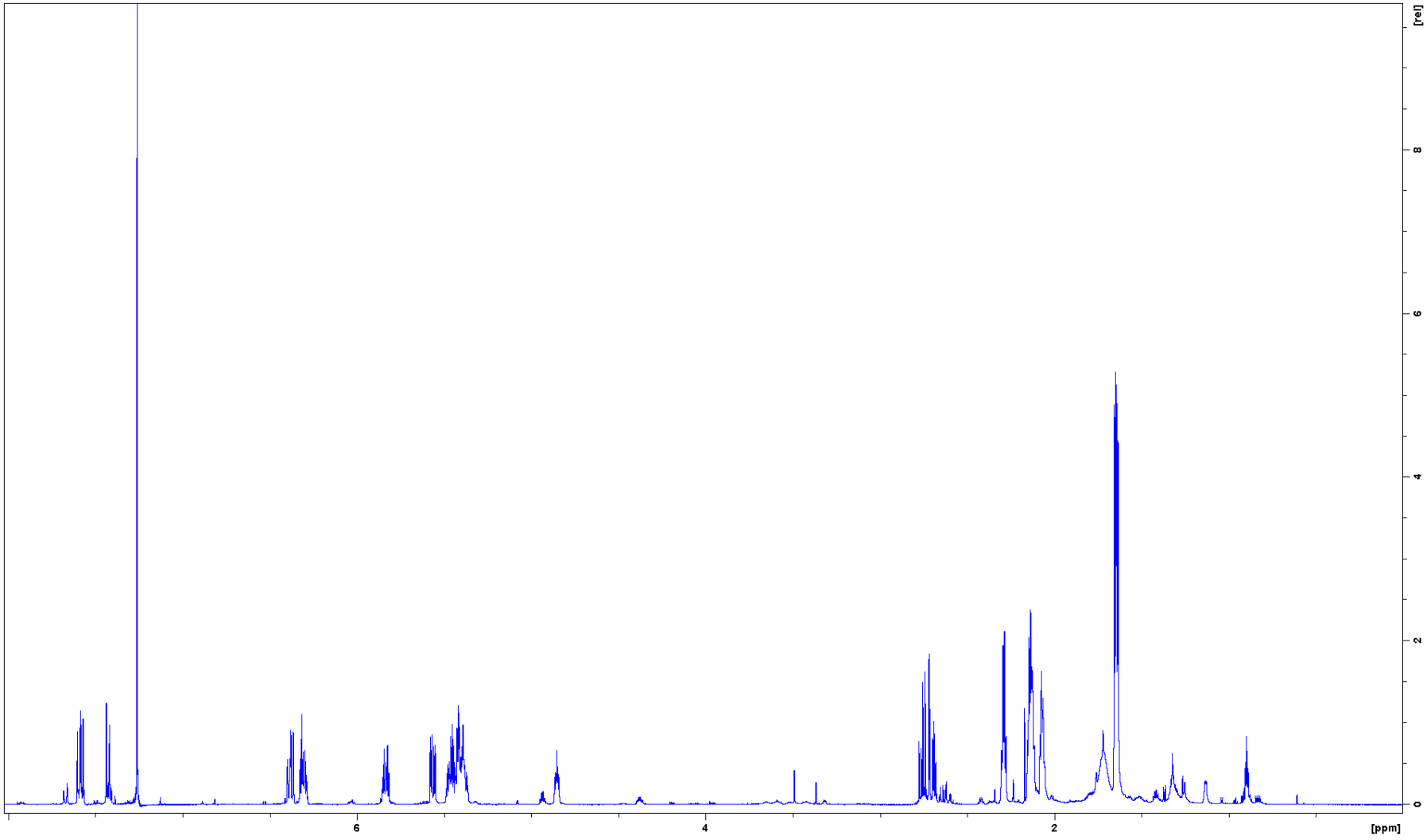


Brasenol B1

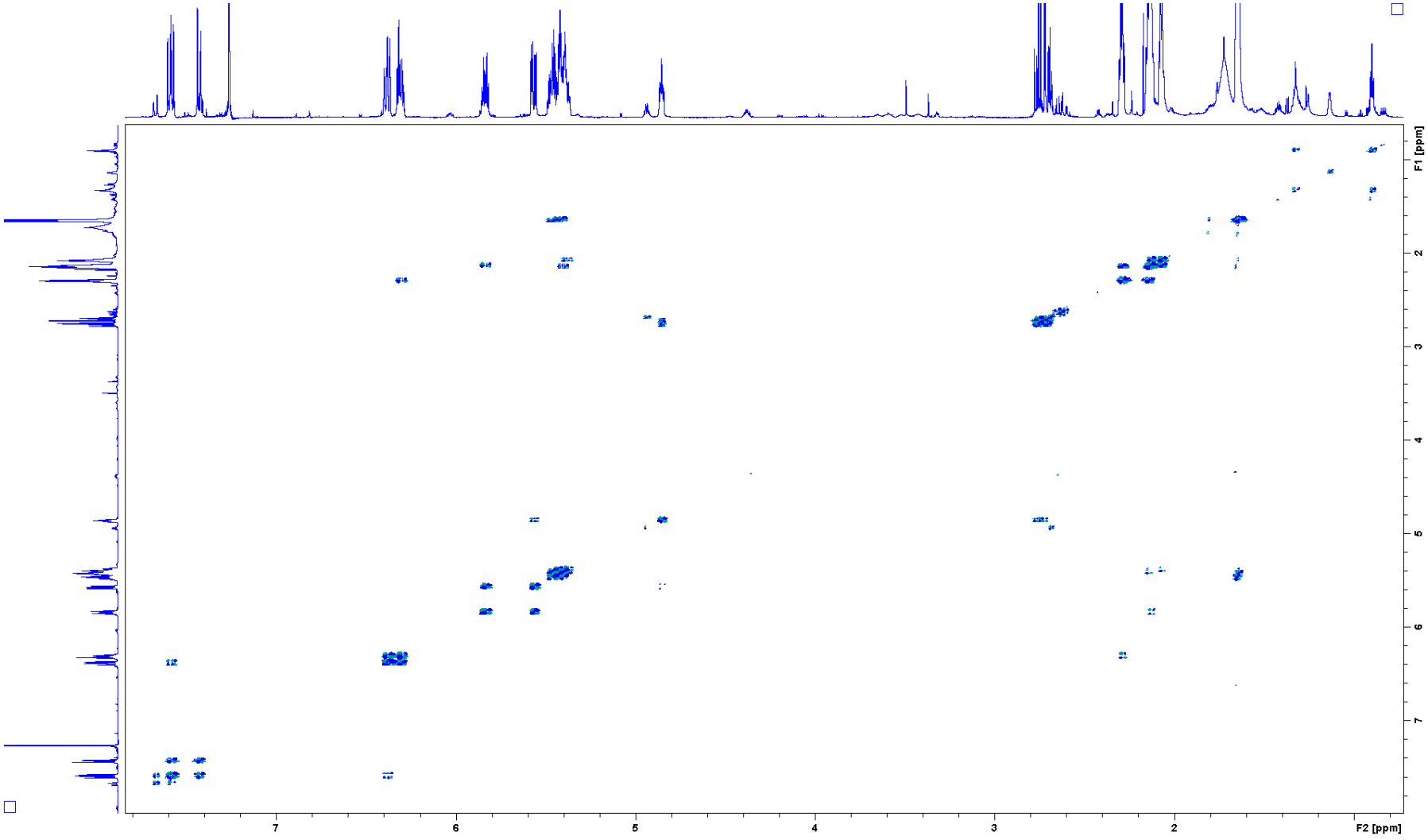
^{1}H-NMR (full range)



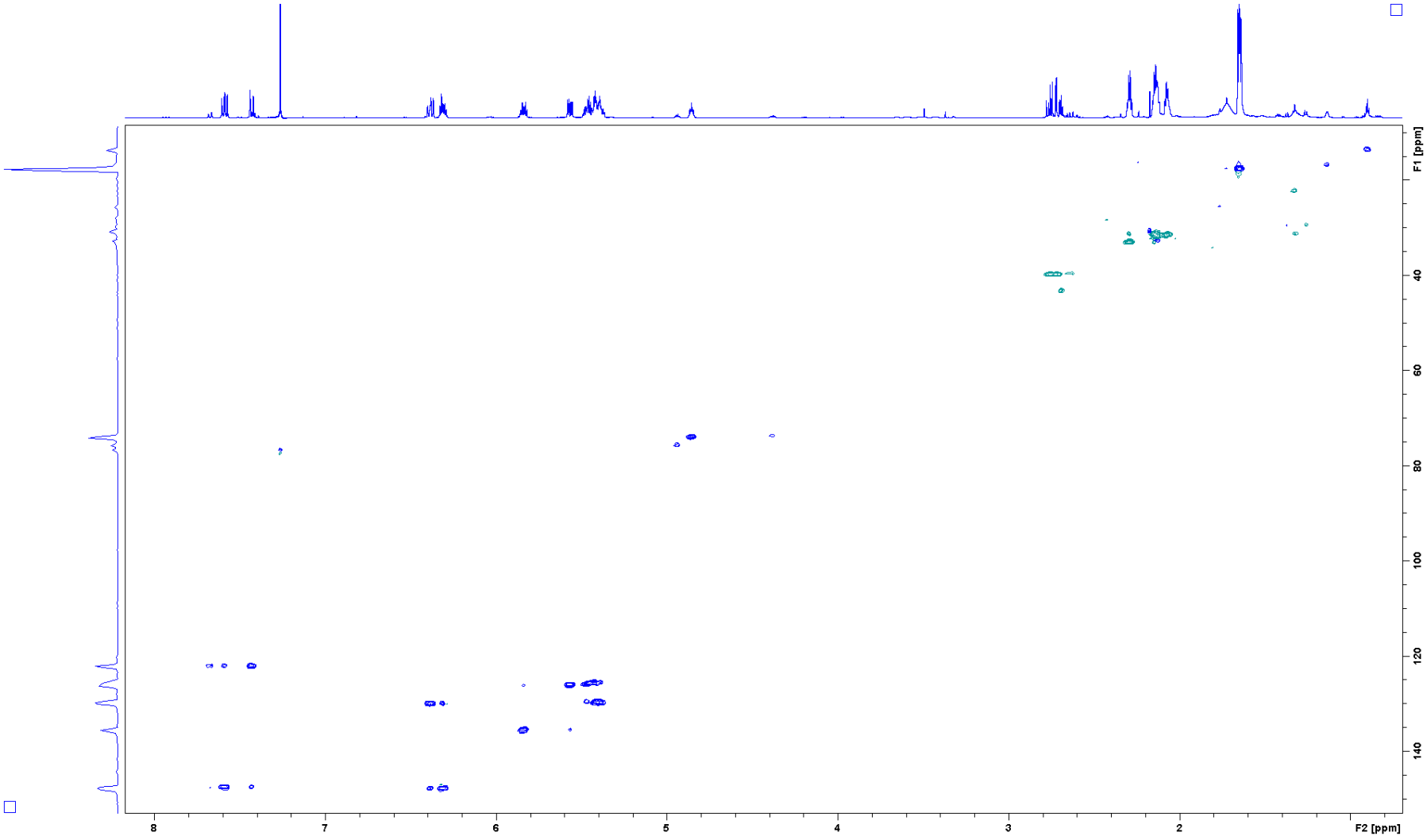
^{1}H-NMR (0-8 Hz)



DQF-COSY

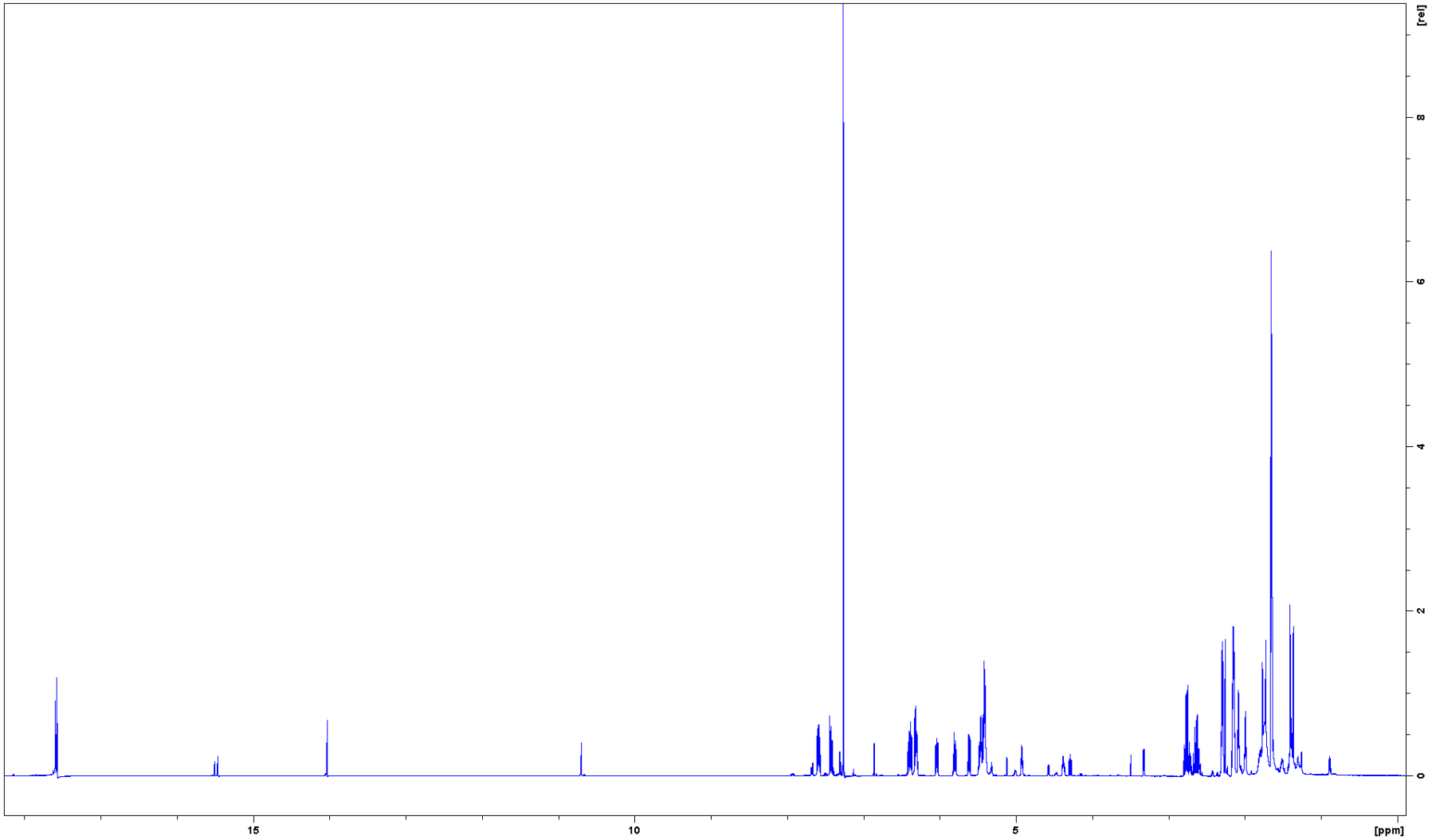


edHSQC

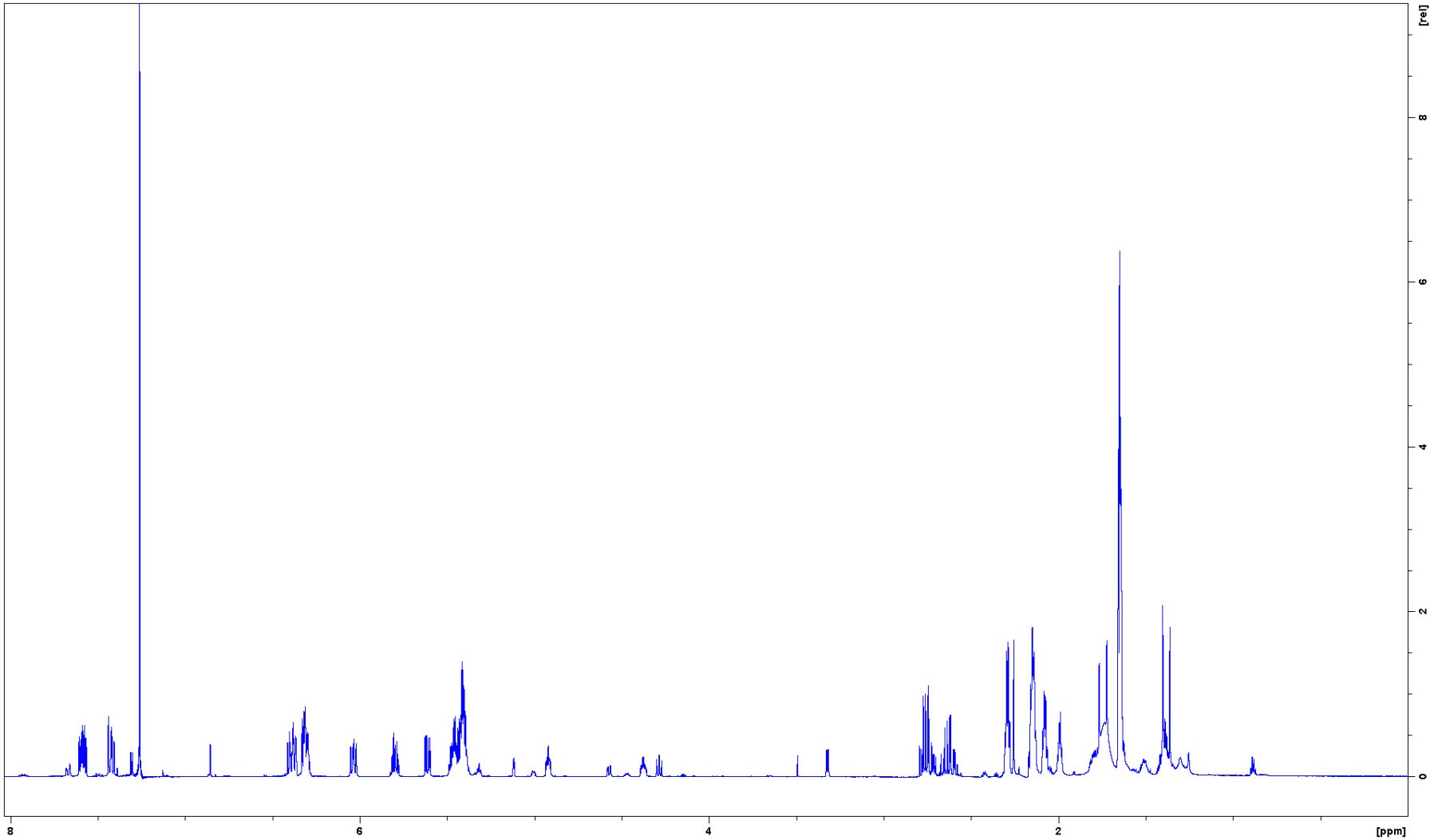


Brasenol B2 + C1 mix

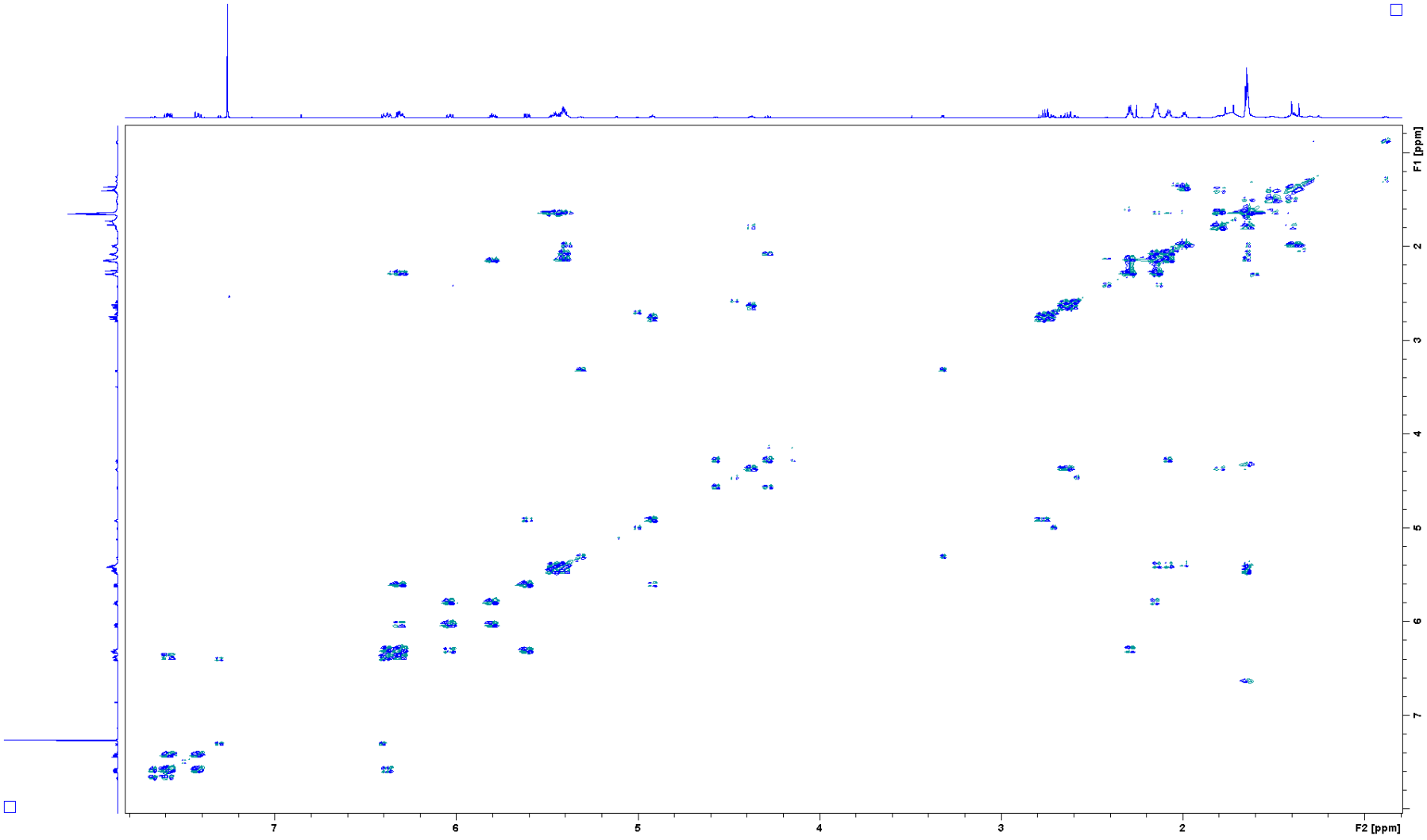
^{1}H-NMR (full range)



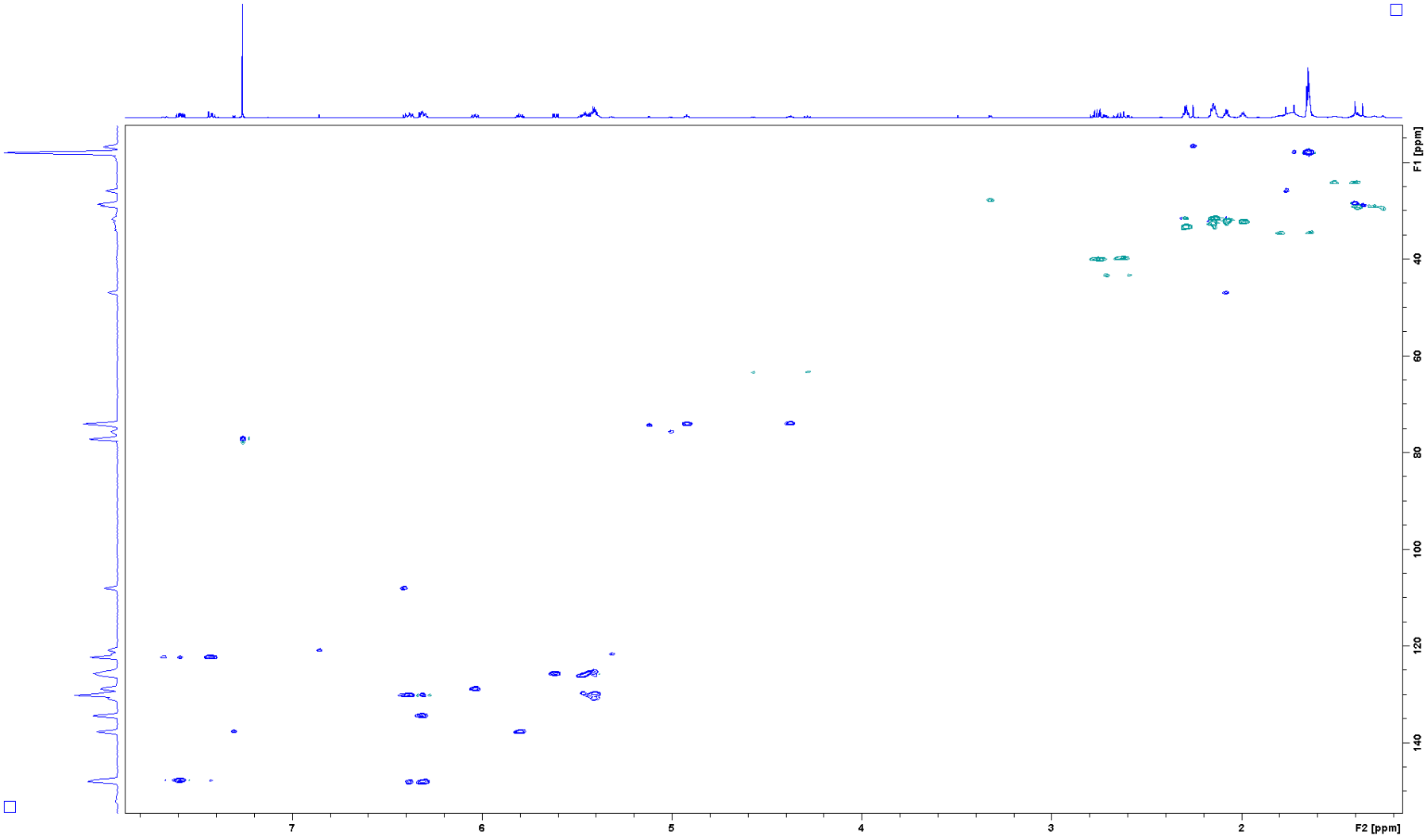
^{1}H-NMR (0-8 Hz)



DQF-COSY



edHSQC



S4. NMR table for brasenol A1, B1, B2 and C1

Table S1. Proton and carbon shifts for brasenol A1, B1, B2, and C1

	Brasenol A1		Brasenol B1		Brasenol B2		Brasenol C1	
#	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C
1	1.63 (m)	17.6	1.64 (m)	17.6	1.64 (m)	17.6	1.64 (m)	17.9
2	5.45 (m)	126.1	5.46 (m)	125.7	5.46 (m)	126.0	5.46 (m)	126.0
3	5.36 (m)	129.1	5.39 (m)	129.7	5.40 (m)	129.9	5.40 (m)	129.9
4	2.1 (m)	31.3	2.13 (m)	31.7	2.14 (m)	31.6	2.14 (m)	31.6
5	2.25 (q, *J* = 7.2 Hz)	33.1	2.29 (q, *J* = 7 Hz)	33.1	2.29 (q, *J* = 7.3 Hz)	33.1	2.29 (q, *J* = 7.3 Hz)	33.4
6	6.27 (m)	147.6	6.31 (m)	147.8	6.31 (m)	147.6	6.31 (m)	148.1
7	6.34 (dd, *J* = 11/14.8 Hz)	129.8	6.37 (dd, *J* = 11/14.8 Hz)	130.0	6.37 (m)	130.0	6.37 (m)	130.2
8	7.54 (dd, *J* = 10.8/15 Hz	147.3	7.58 (dd, *J* = 10.9/15.1 Hz)	147.6	7.58 (m)	147.6	7.58 (m)	147.8
9	7.38 (d, *J* = 15 Hz)	122.0	7.42 (d, *J* = 15 Hz)	122.1	7.42 (d, *J* = 11 Hz)	122.1	7.43 (d, *J* = 11 Hz)	122.35
10	-	186.3	-	-*	-	-*	-	-*
10-OH	17.59 (s)	-	17.59 (s)	-	17.60 (s)	-	17.60 (s)	-
11	-	101.7	-	-*	-	-*	-	-*
12	-	164.7	-	-*	-	-*	-	-*
1'	1.63 (m)	17.6	1.64 (m)	17.6	1.64 (m)	17.6	1.64 (m)	17.9
2'	5.45 (m)	126.1	5.46 (m)	125.7	5.46 (m)	126.0	5.46 (m)	126
3'	5.36 (m)	129.1	5.39 (m)	129.7	5.4 (m)	129.9	5.40 (m)	129.9
4'	2.15 (m)	27.5	2.07 (m)	31.6	1.99 (q, *J* = 6.6 Hz)	32.4	2.07 (q, *J* = 7.3 Hz)	32.1
5'	1.64/1.84 (m)	34.2	2.12 (m)	31.7	1.38 (m)	29.2	2.15 (m)	32.7
6'	4.35 (m)	72.9	5.83 (m)	135.5	1.40/1.50 (m)	24.2	5.79 (m)	137.7
7'	2.55/2.59 (m)	39.5	5.56 (*J* = 15.4/6.5 Hz)	126.1	1.63/1.79 (m)	34.6	6.03 (dd, *J* = 10.5/15.2 Hz)	128.8
8'	-	198.3	4.85 (m)	74.1	4.93 (m)	73.9	6.31 (m)	134.4
9'	/	/	2.72 (m)	39.9	2.63 (m)	39.9	5.61 (dd, 6.4/15.2 Hz)	125.7
10'	/	/	-	-*	-	-*	4.91 (m)	74.1
11'	/	/	/	/	/	/	2.75 (m)	40.1
12'	/	/	/	/	/	/	-	-*

* Quaternary carbons not detected in HSQC.

S5. PCR scheme

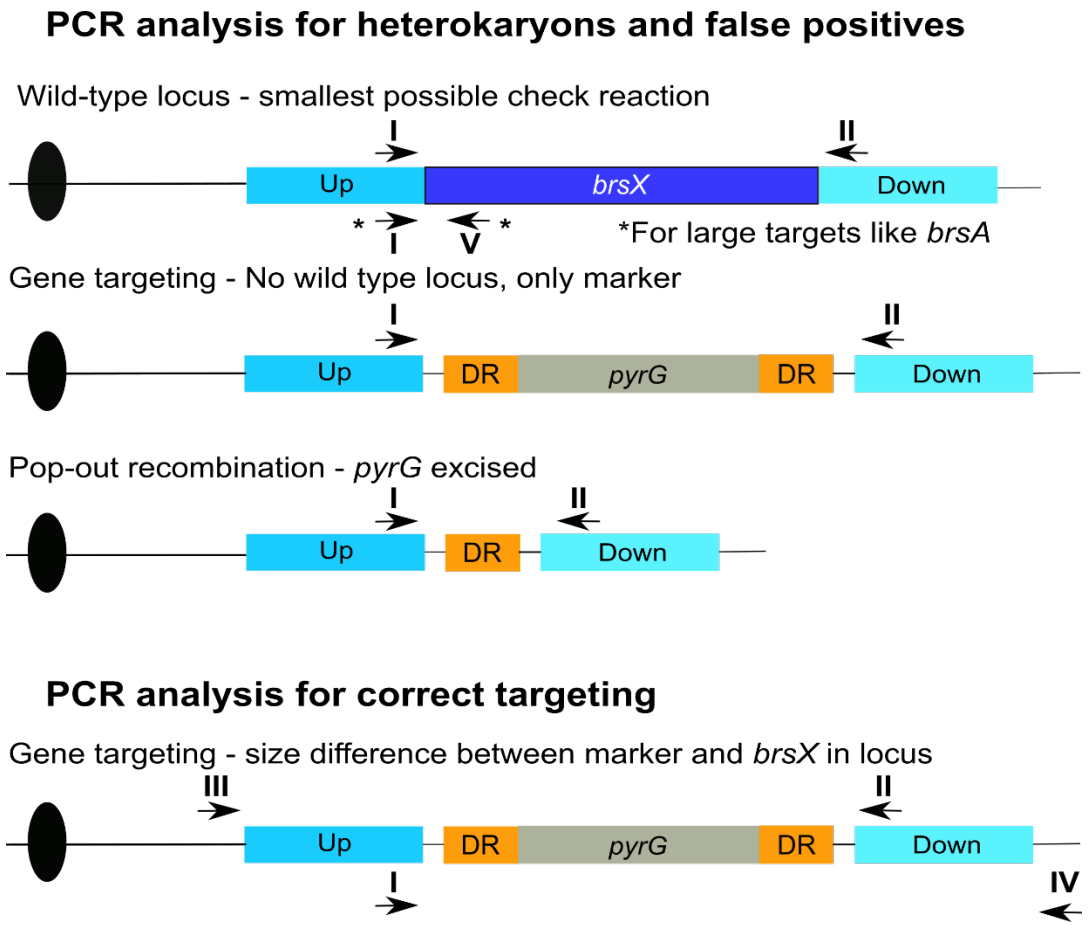


*Figure S2. Diagnostic PCR scheme. Reactions with primer type I and II are intended to show any wild-type (WT) locus bands (*brsX = brsA-D*). In cases, no bands corresponding to WT locus size appear, but only the one of the marker, there is strong indication that the strain is not WT for the locus analysed. This is supported by the reaction with primer types III and II, where one of the primers is unique to the specific locus of integration and not binding within the targeting substrate. The WT locus product may be hard to amplify, when the gene targeted is large, for example in the case of a gene encoding a PKS. Therefore, an internal primer type IV is used together with type I for amplifying a short and simple product revealing any nuclei WT for this locus.*

S6. Primer list

Table S2.Primers used in the study and the reactions

Code[A]	Primer name	Sequence	Purpose & reaction
dUp1	ABRA145702-Dl-Up-FU	GGGTTTAAUGCATGGATCTGTATGCAGTAG	Upstream targeting fragment for deletion of *brsA* – dUp1+dUp2
dUp2	ABRA145702-Dl-Up-RU	GGACTTAAUGGTTGGTGTGGATGGATTGG	Upstream targeting fragment for deletion of *brsA* – dUp1+dUp2
dDw1	ABRA145702-Dl-Dw-FU	GGCATTAAUGGTGTGTACTTCTGTAGTAATGG	Downstream targeting fragment for deletion of *brsA* – dDw1+dDw2
dDw2	ABRA145702-Dl-Dw-RU	GGTCTTAAUCTGTAGGAGTGAGTATAGGGTG	Downstream targeting fragment for deletion of *brsA* – dDw1+dDw2
dUp3	ABRA170790-Up-Dl-FU	GGGTTTAAUGCATGTAGAGAGTGTACAGGG	Upstream targeting fragment for deletion of *brsB* – dUp3+dUp4
dUp4	ABRA170790-Up-Dl-RU	GGACTTAAUCTGGGTGTCGAGTCGATG	Upstream targeting fragment for deletion of *brsB* – dUp3+dUp4
dDw3	ABRA170790-Dl-Dw-FU	GGCATTAAUGGGACCTGATGAGTGATTGAG	Downstream targeting fragment for deletion of *brsB* – dDw3+dDw4
dDw4	ABRA170790-Dl-Dw-RU	GGTCTTAAUGAGTATTCGCTGTGGAGGATG	Downstream targeting fragment for deletion of *brsB* – dDw3+dDw4
dUp5	ABRA39116-Dl-Up-FU	GGGTTTAAUGAGTTGATGTGAACGCAAGAG	Upstream targeting fragment for deletion of *brsC* – dUp5+dUp6
dUp6	ABRA39116-Dl-Up-RU	GGACTTAAUGATGGTTTAGCGATGATATAAGTAGG	Upstream targeting fragment for deletion of *brsC* – dUp5+dUp6
dDw5	ABRA39116-Dl-Dw-FU	GGCATTAAUGGCTATATGGATTTCCTGGGATAG	Downstream targeting fragment for deletion of *brsC* – dDw5+dDw6
dDw6	ABRA39116-Dl-Dw-RU	GGTCTTAAUGAAATCTGAACCGGGTGAAAGG	Downstream targeting fragment for deletion of *brsC* – dDw5+dDw6
dUp7	ABRA39117-Dl-Up-FU	GGGTTTAAUCTCTCGTCCACTCTCAAGAC	Upstream targeting fragment for deletion of *brsD* – dUp7+dUp8
dUp8	ABRA39117-Dl-Up-RU	GGACTTAAUCATCTATTCTACAACCACAACCTC	Upstream targeting fragment for deletion of *brsD* – dUp7+dUp8
dDw7	ABRA39117-Dl-Dw-FU	GGCATTAAUGATGGTCGAGTTGCTCGTTTC	Downstream targeting fragment for deletion of *brsD* – dDw7+dDw8
dDw8	ABRA39117-Dl-Dw-RU	GGTCTTAAUGACGTGACTAACAGATCGTGG	Downstream targeting fragment for deletion of *brsD* – dDw7+dDw8
OexI1	ABRA145702-Oex-FU-Pac	GGGTTTAAUATGGCTGAACCAATCGCG	5´fragment for overexpression of *brsA* – OexI1+ OexI4
OexI2	ABRA145702-Oex-RU-Pac	GGTCTTAAUCTACTGCTTACGCGCGGC	5´fragment for overexpression of *brsA* – OexI1+ OexI4
OexI3	ABRA145702-Oex-int-FU	ATGCTGACCGUGGTCTGG	3´fragment for overexpression of *brsA* – OexI2+ OexI3
OexI4	ABRA145702-Oex-int-RU	ACGGTCAGCAUGCACTGC	3´fragment for overexpression of *brsA* – OexI2+ OexI3
OexII1	ABRA170790-Oex-FU	GGGTTTAAUATGTCTAGCAAGCAGAGAGTCG	Overexpression of *brsB* – OexII1+ OexII2
OexII2	ABRA170790-Oex-RU	GGTCTTAAUTCAAGGCTTCGGCCCAAAGATTTC	Overexpression of *brsB* – OexII1+ OexII2
OexIII1	ABRA145702-Oex1dual-FU	ATCAACAUGGCTGAACCAATCGCG	Overexpression of *brsA* for Pdual– OexIII1+ OexIII2
OexIII2	ABRA145702-Oex1dual-RU	GGGTTTAAUCTGAGCATGATTGATATGTTGGGC	Overexpression of *brsA* for Pdual– OexIII1+ OexIII2
OexIII3	ABRA170790-Oex2dual-FU	ATCGTCAATAAUGTCTAGCAAGCAGAGAGTCG	Overexpression of *brsB* for Pdual– OexIII3+ OexIII4
OexIII4	ABRA170790-Oex2dual-RU	GGTCTTAAUCAGCTAAGCATCGCAGCTG	Overexpression of *brsB* for Pdual– OexIII3+ OexIII4
OexIII5	Pdual-F	ATGTTGAUGTGTGAAGAGATTTAAGGT	Overexpression of PH3/H4 from *A. nidulans* – OexIII5+ OexIII6

OexIII6	Pdual-R	ATTATTGACGAUGAGTTTTGATGGAATTAG	Overexpression of PH3/H4 from *A. nidulans* – OexIII5+ OexIII6
I1	ABRA145702-Chk-Gap-F	GGAGTTGAGAGGGATGAGGATG	Check PCR for wild-type *brsA* nuclei - I1+II1
II1	ABRA145702-Chk-Gap-R	GAGCATGATTGATATGTTGGGC	Check PCR for wild-type *brsA* nuclei - I1+II1
V1	ABRA145702-Chk-Int5'-R	GACGGTGATTGTAGAGGGGATTG	Check PCR for wild-type *brsA* nuclei - I1+V1
III1	ABRA145702-Chk-Up-F	GAGGCTGGGTGTCGAGTC	Check PCR for *brsA* replacement with marker and correct mutant - III1+II1
I2	ABRA170790-Gap-Chk-F	GACTCTCCTGTCACAATAACTCC	Check PCR for *brsB* replacement with marker and correct mutant – I2+II2
III2	ABRA170790-Chk-Up-F	CATCCTCATCCCTCTCAACTC	Check PCR for *brsB* replacement with marker and correct mutant – III2+II2
II2	ABRA170790-Chk-Dw-R	CACCTCTTAACATCTGCTCGC	Check PCR for *brsB* replacement with marker and correct mutant – I2+II2
I3	ABRA39116-ChkGap-F	GAGATCTAGGGTTCGGAGTAGAG	Check PCR for wild-type *brsC* nuclei – I3+II3
IV1	ABRA39116-ChkDw-R	GGAGACTAGTGAACTGTGTGTG	Check PCR for *brsC* replacement with marker and correct mutant – I3+IV1
II3	ABRA39116-ChkGap-R	GTAGATGTAGTGAGTCTGTTAGGG	Check PCR for wild-type *brsC* nuclei – I3+II3
I4	ABRA39117-ChkGap-F	GGGATTAGAGTGAAACAGGGTTG	Check PCR for wild-type *brsD* nuclei – I4+II4
II4	ABRA39117-ChkGap-R	GCGATAGATAGATGTATGTTTGGG	Check PCR for wild-type *brsD* nuclei – I4+II4

[A] Code for use. dUp = for upstream targeting fragment, dDw = for downstream targeting fragment. OexI = for single overexpression of *brsA*. OexII = for single overexpression of *brsB*. OexIII = Overexpression of brsA and brsB unde the control of a dual promoter. I to V, see Supplementary figure XX for explanation of reactions.

RESEARCH ARTICLE

Genes Linked to Production of Secondary Metabolites in *Talaromyces atroroseus* Revealed Using CRISPR-Cas9

Maria Lund Nielsen, Thomas Isbrandt, Kasper Bøwig Rasmussen¤, Ulf Thrane, Jakob Blæsbjerg Hoof, Thomas Ostenfeld Larsen, Uffe Hasbro Mortensen*

Department of Biotechnology and Biomedicine, Technical University of Denmark, Søltofts Plads, Kongens Lyngby, Denmark

¤ Current address: Novo Nordisk A/S, Niels Steensens Vej 1, Gentofte, Denmark
* um@bio.dtu.dk

OPEN ACCESS

Citation: Nielsen ML, Isbrandt T, Rasmussen KB, Thrane U, Hoof JB, Larsen TO, et al. (2017) Genes Linked to Production of Secondary Metabolites in *Talaromyces atroroseus* Revealed Using CRISPR-Cas9. PLoS ONE 12(1): e0169712. doi:10.1371/journal.pone.0169712

Editor: Richard A Wilson, University of Nebraska-Lincoln, UNITED STATES

Received: September 12, 2016

Accepted: December 20, 2016

Published: January 5, 2017

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by a grant from the Novo Nordisk Foundation (http://novonordiskfonden.dk/en). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

The full potential of fungal secondary metabolism has until recently been impeded by the lack of universal genetic tools for most species. However, the emergence of several CRISPR-Cas9-based genome editing systems adapted for several genera of filamentous fungi have now opened the doors for future efforts in discovery of novel natural products and elucidation and engineering of their biosynthetic pathways in fungi where no genetic tools are in place. So far, most studies have focused on demonstrating the performance of CRISPR-Cas9 in various fungal model species, and recently we presented a versatile CRISPR-Cas9 system that can be successfully applied in several diverse *Aspergillus* species. Here we take it one step further and show that our system can be used also in a phylogenetically distinct and largely unexplored species from the genus of *Talaromyces*. Specifically, we exploit CRISPR-Cas9-based genome editing to identify a new gene in *T. atroroseus* responsible for production of polyketide-nonribosomal peptide hybrid products, hence, linking fungal secondary metabolites to their genetic origin in a species where no genetic engineering has previously been performed.

Introduction

Filamentous fungi are known as prolific producers of numerous industrially important enzymes as well as a diverse spectrum of natural products. The latter constitutes an immense reservoir of compounds of biological and medical interest, and many products originating from fungal secondary metabolism are used today in the pharmaceutical industry as *e.g.* antibiotics, anticancer drugs, cholesterol-lowering agents, and immunosuppressive drugs [1]. In addition, some natural products are used commercially as pigments in cosmetics, textiles, paints, and as natural food colorants [2].

The lack of genetic tools available for most fungal species has for many years been the major obstacle for exploring the molecular biology and biochemistry of all but a few model fungi. The enormous increase in sequencing projects over the past years has revealed the

existence of an abundance of uncharacterized and often silent secondary metabolite gene clusters that still awaits investigation [3]. Genetic engineering of the largely unexplored fungal species would thus allow the full study of such gene clusters and could lead to the discovery and characterization of new bioactive compounds.

CRISPR-Cas9-based genetic engineering has recently been implemented in *Aspergillus nidulans* by us [4], and by others in several other species of filamentous fungi such as *Trichoderma reesei* [5], *Neurospora crassa* [6], *Magnaporta oryzae* [7], and *Penicillium chrysogenum* [8]. Our system for *A. nidulans* is based on an AMA1 based vector carrying genes encoding Cas9 and the sgRNA necessary for guiding the Cas9 endonuclease to the desired target sequence and can potentially be used in many fungi with little or no adaptation. In fact, the versatility of this system is demonstrated by the fact that RNA guided mutation was achieved in six different *Aspergillus* species of which one was genetically engineered for the first time [4]. Our CRISPR-Cas9 system may therefore be functional in a wide array of filamentous fungi.

In a recent publication we reported a case of synthetic biology using *A. nidulans* as a host for heterologous gene expression [9]. In this study we successfully exchanged PKS- and NRPS modules between two related PKS-NRPS hybrids to produce the predicted combinations of backbone polyketide-nonribosomal peptide (PK-NRP) products [9]. However, as a surprise the synthetic polyketide-nonribosomal products contained a decalin ring system in the polyketide moiety as well as an unexpected double bond in the amino acid residue side chain, instead of the expected classical cytochalasin structure (Fig 1A and 1B). The structures of these novel derivatives are very similar to ZG-1494α (Fig 1C), an inhibitor of platelet-activating factor acetyltransferase isolated from two species of *Talaromyces*, that is *T. convolutes* and *T. atroroseus* [10,11].

For future synthetic biology efforts, we were therefore interested in identifying the genes and enzymes that are required for production of this scaffold. In *A. nidulans*, we speculate that the unexpected structural features in the synthetic PK-NRPs are a consequence of cross-chemical reactions catalyzed by unknown endogenous enzymes provided by the host. Hence, identification of the genes in *A. nidulans* is not straight forward as the origin of the chemistry is unclear. *T. atroroseus*, which besides ZG-1494α also produces its stereoisomer talaroconvolutin A, and the analogue talaroconvolutin B (Fig 1D), has recently been sequenced [12]. Despite that no genetic tools were available for this species prior to our work, we took advantage of this sequence and set out to identify the genetic basis of ZG-1494α and the related compound talaroconvolutin A in *T. atroroseus* using a bioinformatics approach and our fungal CRISPR-Cas9 technology.

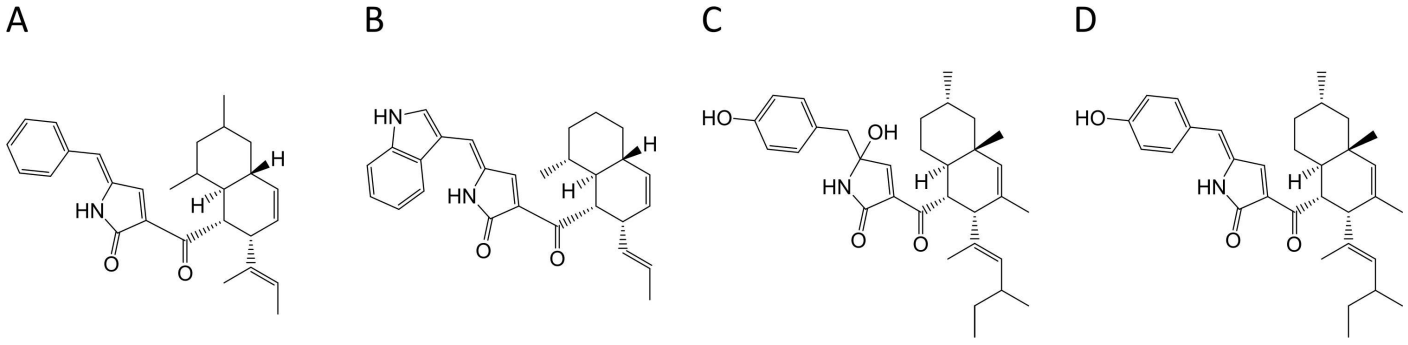


Fig 1. Four structurally similar polyketide-nonribosomal peptide products. Niduclavin (A) and niduporthin (B) are two novel hybrid products produced by heterologous expression of two related PKS-NRPSs in *Aspergillus nidulans* [9], while ZG-1494α/talaroconvolutin B (C) and talaroconvolutin A (D) are produced in *Talaromyces atroroseus*.

doi:10.1371/journal.pone.0169712.g001

Results and Discussion

First we tested whether the CRISPR-Cas9 system that we have previously developed for *Aspergillus* could be used directly in *T. atroroseus*. As a simple test case for Cas9 mediated gene targeting we decided to delete the gene responsible for the green conidia pigment in *T. atroroseus*, which we hypothesized was formed from naphtha-γ-pyrone. To identify this gene we blasted the *A. nidulans*- and *A. niger* naphtha-γ-pyrone synthase genes (*wA* and *albA*, respectively) against the genome sequence of *T. atroroseus*. Amongst the homologous sequences identified in this manner, UA08_00425 was the best match as judged by the size of the ORF and by the high sequence similarities to the corresponding enzymes encoded by *wA* from *A. nidulans* (ID: 62%; 99% query coverage) and *albA* from *A. niger* (ID: 63%; 99% query coverage). UA08_00425 was therefore selected for deletion.

Using our genetic tool box for fungal CRISPR-Cas9 gene editing [4], we constructed a plasmid containing a *hph* selection marker-based gene-targeting substrate designed for deleting UA08_00425, see Fig 2A. Since the efficiency of different protospacers are known to vary substantially [13], three AMA1 based CRISPR-Cas9 vectors encoding Cas9 and one of three different UA08_00425 specific sgRNAs were also constructed (Fig 2A and 2B).

Unlike classical gene targeting where the ends of linear gene-targeting substrates stimulate integration into the target site as they attract the homologous recombination (HR) repair machinery [14]; we exploit that a specific Cas9 induced DNA double strand break (DSB) at the target locus attracts the HR machinery [4]. As a consequence, efficient gene targeting can be achieved by using circular gene-targeting substrates as template for repair of this DNA DSB, hence, minimizing undesirable random integrations mediated by the non-homologous end-joining pathway. Using this strategy for deleting UA08_00425 (Fig 2A and 2C), we investigated

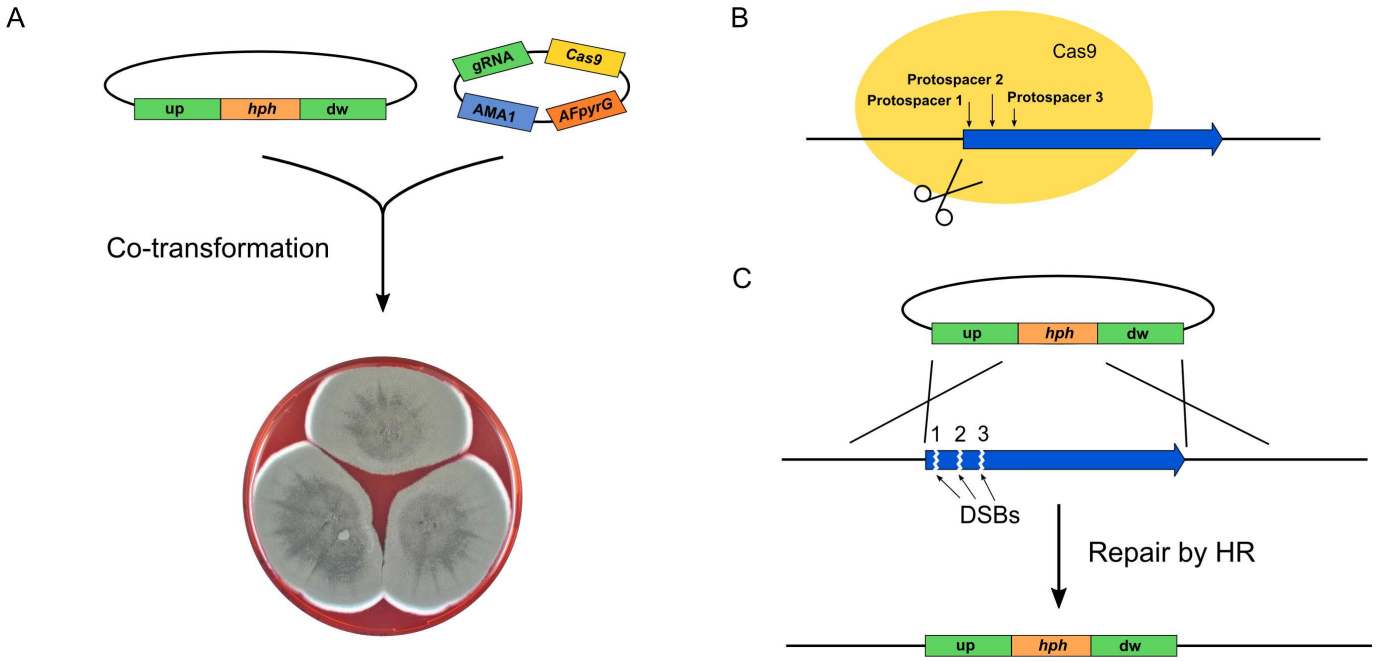


Fig 2. Strategy for CRISPR-Cas9 mediated deletions in *Talaromyces atroroseus*. A) *T. atroroseus* protoplasts were co-transformed with a circular gene-targeting substrate and an AMA1 based CRISPR-Cas9 vector containing Cas9- and sgRNA encoding genes. The transformed protoplasts were plated on medium containing hygromycin, hence, selecting for the gene-targeting substrate only. B) Three different protospacers were individually used to target the Cas9 endonuclease to UA08_00425. The positions are indicated by small vertical arrows. C) Depending on the protospacer contained by Cas9, a specific DNA DSB was produced at either position 1, 2, or 3. Repair of any of these specific DNA DSBs by homologous recombination using the circular gene-targeting substrate as repair template mediate replacement of UA08_00425 with *hph*.

doi:10.1371/journal.pone.0169712.g002

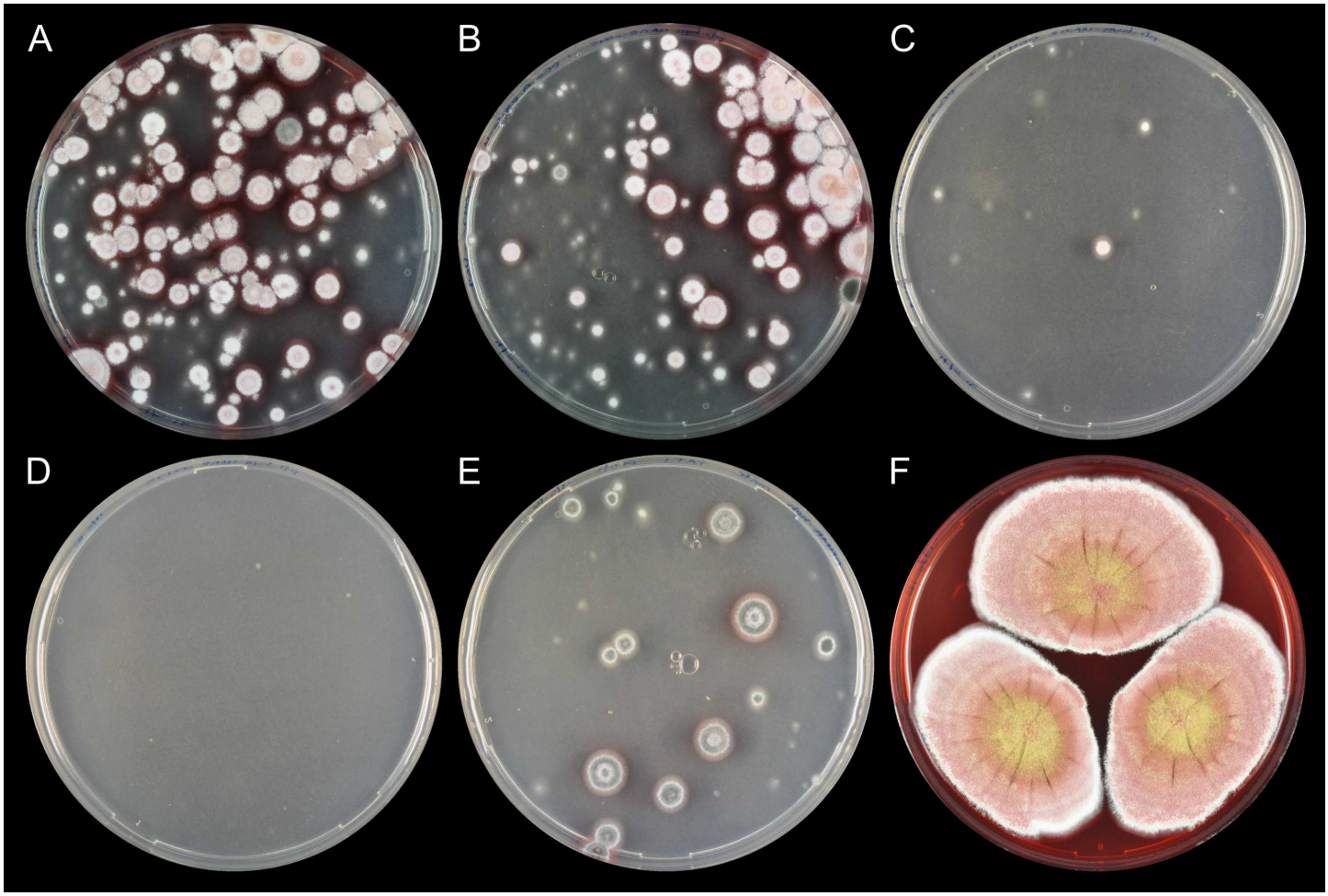


Fig 3. Deletion of the green pigment gene UA08_00425 in *Talaromyces atroroseus* using CRISPR-Cas9. A-C) Plates resulting from co-transformation of pD-hyg-UA08_00425 and CRISPR-Cas9 vectors carrying three different protospacers, protospacer 1–3, respectively. D) *T. atroroseus* transformed with gene-targeting plasmid pD-hyg-UA08_00425 (pFC574) in the absence of a CRISPR-Cas9 vector. E) *T. atroroseus* transformed with an AMA1-based plasmid containing the hygromycin resistance marker *hph*. F) Three-point inoculation of UA08_00425Δ *T. atroroseus* growing on CYA.

doi:10.1371/journal.pone.0169712.g003

the Cas9 mediated gene targeting efficiency in *T. atroroseus* by co-transforming the circular vector containing the UA08_00425 gene deletion sequence with each of the three UA08_00425 specific CRISPR-Cas9 vectors. AMA1 based plasmids are readily lost in the absence of selection pressure [15] and this is also the case for our AMA1 based CRISPR plasmids (see S1 Fig). To reduce the risk of undesired off-target effects, we therefore selected for the gene-targeting substrate only, and not for the *cas9* expressing AMA1 based plasmid, hence, confining *cas9* expression to the early stages of colony development.

In two independent trials, two of the three co-transformation experiments generated numerous colonies on solid selective medium after approximately one week, whereas the remaining co-transformation produced only a few colonies (Fig 3A–3C and S2 Fig). Importantly, on all three transformation plates, close to all colonies formed white conidia spores in agreement with UA08_00425 encoding the naphtha-γ-pyrone PKS. In contrast, in the absence of Cas9, either no colonies or only green colonies were observed (Fig 3D and S2 Fig). More importantly, the results strongly indicate that Cas9 has efficiently stimulated gene deletion of UA08_00425 in these experiments. This conclusion is substantiated by the results of two

control experiments. Firstly, no transformants were obtained when the circular gene-targeting substrate was transformed alone into *T. atroroseus* (Fig 3D) indicating that the specific Cas9 induced DNA DSB is required for integrating information from the circular gene-targeting substrate into the UA08_00425 locus. Secondly, no white transformants were obtained with pFC574 carrying only the *hph* gene. This control experiment shows that white conidia spores are not due to the presence of hygromycin *per se* (Fig 3E). Finally, we note that integration efficiencies are approximately 10-fold more efficient with protospacer 1 and 2 as compared to the efficiency obtained with protospacer 3.

Next, we streak purified six white transformants on solid medium without hygromycin selection. In all cases, colonies remained solid white showing that the white phenotype of transformants could be stably propagated as expected from a strain containing a permanent gene deletion (Fig 3F). In agreement with this conclusion, we confirmed that UA08_00425 was eliminated in all six purified strains by tissue-PCR (S3 Fig and S5 Fig). Since deletion of UA08_00425 results in white conidia spores, we have named this gene *albA* (S1 Appendix).

The efficient CRISPR-Cas9 mediated deletion of *albA* in *T. atroroseus* prompted us to identify the genetic origin of ZG-1494α and its derivative talaroconvolutin A (Fig 1).The structures of these compounds appear to be fusions of highly reduced polyketide moieties to tyrosine residues, similar to what is seen for example in cytochalasins and chaetoglobosins [16]. Therefore, we suspected that a homolog of *ccsA*, the PKS-NRPS-encoding gene linked to cytochalasin production in *A. clavatus* [17], is responsible for the biosynthesis of a common backbone for these compounds. In support of this view, the nitrogen-containing tetramic acid moiety present in both compounds is a common structural feature for several known PKS-NRPS products such as preaspyridone, pretenellin A, prepseurotin, as well as niduclavin, and niduporthin [9,18–20]. Due to the structural resemblance between the niduclavin backbone and the backbone of the talaroconvolutins/ZG1494α, we blasted the *ccsA* gene against the *T. atroroseus* genome, to identify PKS-NRPS-encoding genes in *T. atroroseus*. Based on this analysis we selected the gene with the highest sequence identity, UA08_04451 (ID: 46.0%, 96% query coverage) for deletion.

A gene-targeting substrate for deletion of UA08_04451 was constructed and co-transformed with a CRISPR-Cas9 plasmid carrying a sgRNA targeting UA08_04451. After approximately one week, green colonies appeared on solid hygromycin selection medium (Fig 4A). Importantly, no transformants appeared when the gene-targeting substrate was transformed

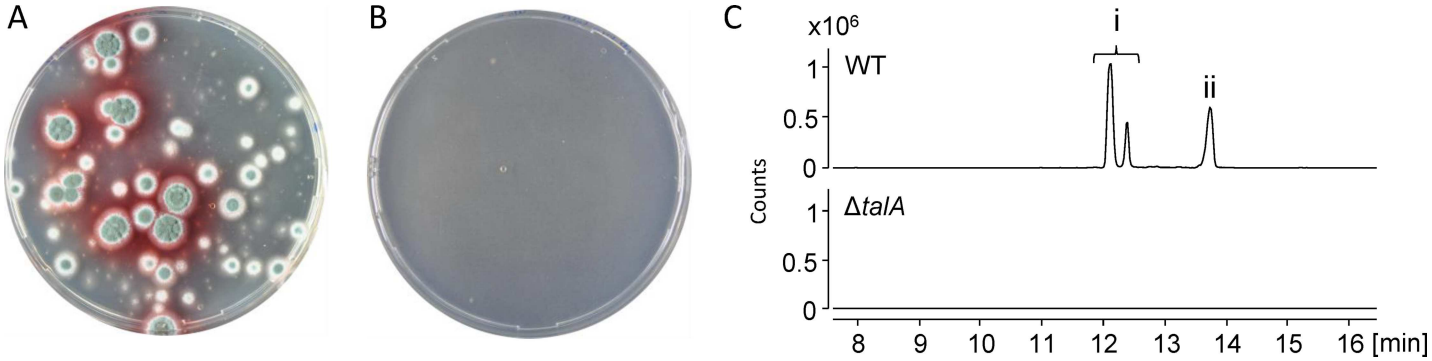


Fig 4. Deletion of UA08_04451 (*talA*), encoding a hybrid PKS-NRPS, in *Talaromyces atroroseus* using CRISPR-Cas9. A) Co-transformation of gene-targeting plasmid pD-hyg-talA with a CRISPR-Cas9 vector. B) *T. atroroseus* transformed with pD-hyg-talA in the absence of a CRISPR-Cas9 vector. C) UHPLC-HRMS analysis of the extracts of wild type (WT) *T. atroroseus* along with a *T. atroroseus talA* deletion strain. Shown are EIC @ 528.3084 (±0.0100) for ZG-1494α (i), and EIC @ 510.2979 (±0.0100) for talaroconvolutin A (ii). We suspect the two peaks in (i) to represent both ZG-1494α and talaroconvolutin B. Note that EICs are based on $[M+Na]^+$ adducts.

doi:10.1371/journal.pone.0169712.g004

without a CRISPR-Cas9 plasmid strongly indicating that formation of the transformants required Cas9 activity (Fig 4B). Eight colonies were streak purified on solid CYA medium supplemented with hygromycin and were subsequently analyzed by tissue PCR. The PCR results confirmed the deletion of UA08_04451 for at least seven out of the eight streak purified candidates (S4 Fig and S5 Fig). The seven UA08_04451 deletion strains were analyzed by UHPLC-HRMS and in all cases production of both talaroconvolutin A and ZG-1494α was abolished (Fig 4C). Together these results strongly indicate that talaroconvolutin A and ZG-1494α are formed from a common PK-NRP backbone synthesized by a PKS-NRPS fusion enzyme encoded by UA08_04451, and we have therefore named this gene *talA* (S2 Appendix).

Conclusions

In this study we have used CRISPR-Cas9 technology to genetically engineer *T. atroroseus* and used it to explore the secondary metabolism of this fungus. Specifically, we have identified a novel gene encoding a hybrid PKS-NRPS, which is responsible for production of medically relevant ZG-1494α. To the best of our knowledge, this represents the first example of reverse genetic engineering of a *Talaromyces* species. Importantly, the fact that our CRISPR-Cas9 system, which we have originally developed for gene editing of *Aspergillus* species, can be used without any modifications to engineer a phylogenetically distinct species, raises the possibility that it can also be used directly in a wide range of other fungal species.

Materials and Methods

Strains, genomic DNA, and media

T. atroroseus strain IBT 11181 was obtained from the IBT Culture Collection at Department of Biotechnology and Biomedicine at Technical University of Denmark. It is also deposited in the CBS collection at CBS-KNAW, the Netherlands, as CBS 123796 and CBS 238.95. The *T. atroroseus* genome sequence has been deposited at DDBJ/ENA/GenBank under the accession LFMY00000000. The version described in this paper is version LFMY01000000. DNA sequences of *T. atroroseus* genes *albA* and *talA* are presented in S1 Appendix and S2 Appendix, respectively. Genomic DNA (gDNA) from *T. atroroseus* was extracted using the FastDNA™ SPIN Kit for Soil DNA extraction (MP Biomedicals, USA), and *T. atroroseus* gDNA was used as PCR template for amplification of the up- and downstream fragments for deletion of *talA* (UA08_04451) and the green pigment gene (UA08_00425). *T. atroroseus* was cultivated in liquid- and on solid CYA medium (Czapek yeast autolysate) supplemented with 300 μg/ml hygromycin B (Hygrogold, Invivogen) when needed. *Escherichia coli* strain DH5α was used for plasmid propagation.

Vector construction

PCR fragments were amplified using the PfuX7 polymerase [21] with primers purchased from Integrated DNA Technology, Belgium (S1 Table). Construction of vectors was carried out by Uracil-Specific Excision Reagent (USER) fusion of PCR fragment into compatible plasmids [22]. The deletion plasmids pD-hyg-talA and pD-hyg-albA were constructed by amplification of approximately 2-kb up- and downstream fragments followed by cloning into two distinct *Pac*I/*Nt.Bbv*CI USER cassettes located on each side of the hygromycin resistance gene. The sgRNA was introduced into the CRISPR-Cas9 vector pFC330 via the tails of two primers as described by Nødvig et al. [4]. Plasmids were purified using the GenElute™ Plasmid Miniprep Kit (Sigma-Aldrich), and verified by restriction analysis. A list of all plasmids from this study is presented in S2 Table. Deletions were achieved using the CRISPR-Cas9 system described by

Nødvig *et al.* [4]. A circular deletion plasmid (gene-targeting substrate) was co-transformed with an AMA1-based CRISPR-Cas9 vector containing the guide RNA and the *Streptococcus pyogenes cas9* gene codon optimized for *A. niger*. The CRISPR-Cas9 vector also contained the *pyrG* auxotrophic marker; however, only the deletion plasmid, containing the hygromycin resistance gene, was selected for during transformation.

Protoplastation and transformation

Protoplastation of *T. atroroseus* was achieved using protocols described previously for *A. nidulans* [23,24]. For transformation, 2.5–3 μg DNA of the deletion plasmid and 2.5–3 μg DNA of the CRISPR-Cas9 vector were mixed with 100 μl protoplasts. 100 μl of a solution of 40% PEG in 1 M sorbitol, 50 mM Tris, 10 mM $CaCl_2$, pH 7.5 was added and the sample was incubated on ice for 15 min. Another 500 μl of the PEG solution was added followed by incubation at room temperature for another 15 min. The mixture was then added to 8 ml molten soft (0.8% agar) CYA medium supplemented with 1 M sorbitol and spread on solid CYA plates supplemented with 1 M sorbitol (2% agar). The plates were incubated O/N at 30˚C, and the next day overlaid with 8 ml soft CYA medium supplemented with 300 μg/ml hygromycin. The plates were incubated at 30˚C until transformants appeared on the transformation plates (approximately 1 week). Transformants were re-streaked on CYA plates containing the same antibiotic concentration. Tissue-PCR as described by Nødvig *et al.* [4] was used for strain validation (see S3 Fig and S4 Fig). Two sets of primers were used to validate the deletions of *talA* and the green pigment gene. In one reaction the reverse primer would bind in the promoter of the marker (P*gpdA*) while the forward primer would bind outside the upstream targeting sequence. In another reaction designed to check for negatives or possible heterokaryons the forward primer would bind in the upstream targeting sequence while the reverse primer would bind inside the gene.

Chemical analysis of *T. atroroseus* strains

Validated *T. atroroseus* strains were grown for 7 days on CYA plates and plug extractions were performed as described by Smedsgaard [25] with the exception that secondary metabolites were extracted with 3:1 ethylacetate:isopropanol containing 1% formic acid. Ultra-high Performance Liquid Chromatography-High Resolution Mass Spectrometry (UHPLC-HRMS) was performed on an Agilent Infinity 1290 UHPLC system (Agilent Technologies, Santa Clara, CA, USA) equipped with a diode array detector. Separation was obtained on an Agilent Poroshell 120 phenyl-hexyl column (2.1 × 250 mm, 2.7 μm) with a linear gradient consisting of water (A) and acetonitrile (B) both buffered with 20 mM formic acid, starting at 10% B and increased to 100% in 15 min where it was held for 2 min, returned to 10% in 0.1 min and remaining for 3 min (0.35 mL/min, 60˚C). An injection volume of 1 μL was used. MS detection was performed in positive detection on an Agilent 6545 QTOF MS equipped with Agilent Dual Jet Stream electrospray ion source with a drying gas temperature of 250˚C, gas flow of 8 L/min, sheath gas temperature of 300˚C and flow of 12 L/min. Capillary voltage was set to 4000 V and nozzle voltage to 500 V. Mass spectra were recorded at 10, 20 and 40 eV as centroid data for *m/z* 85–1700 in MS mode and *m/z* 30–1700 in MS/MS mode, with an acquisition rate of 10 spectra/s. Lock mass solution in 70:30 methanol:water was infused in the second sprayer using an extra LC pump at a flow of 15 μL/min using a 1:100 splitter. The solution contained 1 μM tributylamine (Sigma-Aldrich) and 10 μM Hexakis(2,2,3,3-tetrafluoropropoxy)phosphazene (Apollo Scientific Ltd., Cheshire, UK) as lock masses. The $[M + H]^+$ ions (*m/z* 186.2216 and 922.0098 respectively) of both compounds was used. Extracted ion chromatograms were used to evaluate the production of ZG-1494α and talaroconvolutin A.

Supporting Information

S1 Appendix. DNA sequence of *albA* including 3 kb up- and downstream sequences.
(DOCX)

S2 Appendix. DNA sequence of *talA* including 3 kb up- and downstream sequences.
(DOCX)

S1 Fig. The stability of AMA1 plasmids in *Talaromyces atroroseus*.
(DOCX)

S2 Fig. Deletion of the green pigment UA08_00425 (*albA*) in *T. atroroseus*. Second independent trial.
(DOCX)

S3 Fig. Tissue PCR analysis for verification of *albA* deletion.
(DOCX)

S4 Fig. Tissue PCR analysis for verification of *talA* deletion.
(DOCX)

S5 Fig. Tissue PCR analysis for verification of primer functionality.
(DOCX)

S1 Table. List of primers.
(DOCX)

S2 Table. List of plasmids.
(DOCX)

Acknowledgments

This work was supported by a grant from the Novo Nordisk Foundation (NNF). We thank Dr. Christina Spuur Nødvig for valuable input regarding the CRISPR-Cas9 system and Martin Engelhard Kogle for technical assistance.

Author Contributions

Conceptualization: MLN UT JBH TOL UHM.

Formal analysis: MLN TI KBR.

Funding acquisition: TOL.

Investigation: MLN TI JBH.

Writing – original draft: MLN.

Writing – review & editing: MLN TI UT JBH TOL UHM.

References

1. Newman DJ, Cragg GM. Natural products as sources of new drugs from 1981 to 2014. J Nat Prod. 2016; 79: 629–661. doi: 10.1021/acs.jnatprod.5b01055 PMID: 26852623
2. Tuli HS, Chaudhary P, Beniwal V, Sharma AK. Microbial pigments as natural color sources: current trends and future perspectives. J Food Sci Technol. 2015; 52: 4669–4678. doi: 10.1007/s13197-014-1601-6 PMID: 26243889

3. Brakhage AA. Regulation of fungal secondary metabolism. Nat Rev Microbiol. 2013; 11: 21–32. doi: 10.1038/nrmicro2916 PMID: 23178386
4. Nødvig CS, Nielsen JB, Kogle ME, Mortensen UH. A CRISPR-Cas9 system for genetic engineering of filamentous fungi. PLoS One. 2015; 10: e0133085. doi: 10.1371/journal.pone.0133085 PMID: 26177455
5. Liu R, Chen L, Jiang Y, Zhou Z, Zou G. Efficient genome editing in filamentous fungus Trichoderma reesei using the CRISPR/Cas9 system. Cell Discov. 2015; 1: 15007. doi: 10.1038/celldisc.2015.7 PMID: 27462408
6. Matsu-ura T, Baek M, Kwon J, Hong C. Efficient gene editing in Neurospora crassa with CRISPR technology. Fungal Biol Biotechnol. 2015; 2: 4.
7. Arazoe T, Miyoshi K, Yamato T, Ogawa T, Ohsato S, Arie T, et al. Tailor-made CRISPR/Cas system for highly efficient targeted gene replacement in the rice blast fungus. Biotechnol Bioeng. 2015; 112: 2543–2549. doi: 10.1002/bit.25662 PMID: 26039904
8. Pohl C, Kiel JAKW, Driessen AJM, Bovenberg RAL, Nygård Y. CRISPR/Cas9 based genome editing of Penicillium chrysogenum. ACS Synth Biol. 2016; 5: 754–764. doi: 10.1021/acssynbio.6b00082 PMID: 27072635
9. Nielsen ML, Isbrandt T, Petersen LM, Mortensen UH, Andersen MR, Hoof JB, et al. Linker flexibility facilitates module exchange in fungal hybrid PKS-NRPS engineering. PLoS One. 2016; 11: e0161199. doi: 10.1371/journal.pone.0161199 PMID: 27551732
10. West RR, Van Ness J, Varming AM, Rassing B, Biggs S, Gasper S, et al. ZG-1494 alpha, a novel platelet-activating factor acetyltransferase inhibitor from Penicilium rubrum, isolation, structure elucidation and biological activity. J Antibiot (Tokyo). 1996; 49: 967–973.
11. Suzuki S, Hosoe T, Nozawa K, Kawai KI, Yaguchi T, Udagawa SI. Antifungal substances against pathogenic fungi, talaroconvolutins, from Talaromyces convolutus. J Nat Prod. 2000; 63: 768–772. PMID: 10869198
12. Bøwig K. Talaromyces atroroseus—Genome sequencing, Monascus pigments and azaphilone gene cluster evolution (PhD thesis). 2015; Technical University of Denmark, Kgs. Lyngby, DK.
13. Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, et al. Rational design of highly active sgRNAs for CRISPR-Cas9–mediated gene inactivation. Nat Biotechnol. 2014; 32: 1262–7. doi: 10.1038/nbt.3026 PMID: 25184501
14. Rothstein RJ. One-step gene disruption in yeast. Methods Enzymol. 1983; 101: 202–211. PMID: 6310324
15. Gems D, Johnstone IL, Clutterbuck AJ. An autonomously replicating plasmid transforms Aspergillus nidulans at high frequency. Gene. 1991; 98: 61–67. PMID: 2013411
16. Scherlach K, Boettger D, Remme N, Hertweck C. The chemistry and biology of cytochalasans. Nat Prod Rep. 2010; 27: 869–86. doi: 10.1039/b903913a PMID: 20411198
17. Qiao K, Chooi YH, Tang Y. Identification and engineering of the cytochalasin gene cluster from Aspergillus clavatus NRRL 1. Metab Eng. 2011; 13: 723–732. doi: 10.1016/j.ymben.2011.09.008 PMID: 21983160
18. Eley KL, Halo LM, Song Z, Powles H, Cox RJ, Bailey AM, et al. Biosynthesis of the 2-pyridone tenellin in the insect pathogenic fungus Beauveria bassiana. Chembiochem. 2007; 8: 289–97. doi: 10.1002/cbic.200600398 PMID: 17216664
19. Xu W, Cai X, Jung ME, Tang Y. Analysis of intact and dissected fungal polyketide synthase-nonribosomal peptide synthetase in vitro and in Saccharomyces cerevisiae. J Am Chem Soc. 2010; 132: 13604–13607. doi: 10.1021/ja107084d PMID: 20828130
20. Kakule TB, Lin Z, Schmidt EW. Combinatorialization of fungal polyketide synthase–peptide synthetase hybrid proteins. J Am Chem Soc. 2014; 136: 17882–17890. doi: 10.1021/ja511087p PMID: 25436464
21. Nørholm MHH. A mutant Pfu DNA polymerase designed for advanced uracil-excision DNA engineering. BMC Biotechnol. 2010; 10: 21. doi: 10.1186/1472-6750-10-21 PMID: 20233396
22. Geu-Flores F, Nour-Eldin HH, Nielsen MT, Halkier BA. USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. Nucleic Acids Res. 2007; 35: e55. doi: 10.1093/nar/gkm106 PMID: 17389646
23. Tilburn J, Scazzocchio C, Taylor GG, Zabicky-Zissman JH, Lockington RA, Davies RW. Transformation by integration in Aspergillus nidulans. Gene. 1983; 26: 205–221. PMID: 6368319
24. Nielsen ML, Albertsen L, Lettier G, Nielsen JB, Mortensen UH. Efficient PCR-based gene targeting with a recyclable marker for Aspergillus nidulans. Fungal Genet Biol. 2006; 43: 54–64. doi: 10.1016/j.fgb.2005.09.005 PMID: 16289954
25. Smedsgaard J. Micro-scale extraction procedure for standardized screening of fungal metabolite production in cultures. J Chromatogr A. 1997; 760: 264–270. PMID: 9062989

SUPPORTING INFORMATION

Genes Linked to Production of Secondary Metabolites in *Talaromyces atroroseus* Revealed Using CRISPR-Cas9

Maria Lund Nielsen, Thomas Isbrandt, Kasper Bøwig Rasmussen, Ulf Thrane, Jakob Blæsbjerg Hoof, Thomas Ostenfeld Larsen, Uffe Hasbro Mortensen

Department of Biotechnology and Biomedicine, Technical University of Denmark, Søltofts Plads, Kongens Lyngby, Denmark

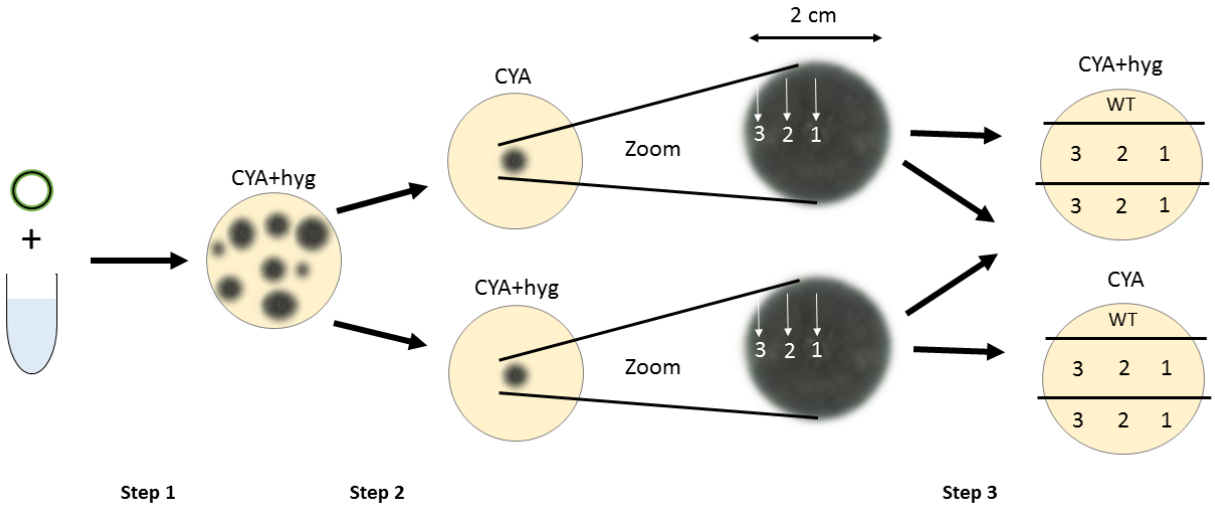
S1 Appendix. DNA sequence of *albA* including 3 kb up- and downstream sequences. The coding sequence is highlighted in yellow, while predicted introns are marked in green.

CATCCTCTTATCGCGGAACAAGCGTCCCATCCACCATGACCCGTCGCCAGCGTTCAGGCATTCTATAAGATCCCCCTCCACAAACCCCAGATCTCTTTTCGATTCCCCTCCCCATGAATATACTGC
TCGACACCAGCATGGAAATTTGGTTGGCAGCGACGGCGGACCGGGCGCCATGGCGAGCGCTTTAAAAGACTCTTGTGAAACGATATGAAGAAAACCCGGCGCTGCAGTGCCAGCTTGAATT
GGCACTGCTCCAAAATTCGAGTATCTGCGACCCAGATTACAGGTCACTAAAGGTAATTGGAAACGAGAGTAAGCGTTCGTTGCTGTCGTGGTTTGGGTACAAGTGAGCGACGGAGCGGATG
TGATAGCTGAGGAAGGGCGGGAATACAATAACAAAGATAATAATGAGCGACCTTCTAGGAGAAAATGACAGTCCGCACACAGTTCATCTGCTGTTCTGTAAGCGACCAGCAGACAGACCAG
AAATAAAACGTTGAGGCCAGACTCATGCGCCAGGCCTTGGTTAGGTAGCGATGCGAAGTCGGTGAGGAGGAGAGCGAACGAGCGTTGCCGTGTTTGGCTGGATGTTTGTGCCCGGTGTTTA
CGCTGCACCTGTTTGGGTTTAGCGCGAGCTGTCTATCGATAAGCTTAACCTCGACGTAGCTACCTGCACTTCCTTTCTCACAGCATGTTTTTTCGAACTCTAGTACTGATGTCGAGCTTCTCTACG
AGCAATTAATAGTAGTCAGGTTTTGTCAATCTGCGCTGCCCTGAATACATCTTGTCTGTACCTTGGGAGTAGGCTGCGGTCAATCAAGTCACTCTACACTGACCACAGTTTCCAACAAGTAAAG
GATTGGGGAAAGCGAGCATCGAATTGGACAGCAACTGGACCAGTATCTTCGTATCTTTCGCGGCACAACTCCTATTTGCACGACAAAAAGGTTGTCGCTCCATTGTTGGTCGTGGTTGTCGAC
GGATATTACTGAGCTCGGATTCACGTCGAAACCAGTCACTACGTTCCGATTTTCGCTATTAATGAGCTTTTAGGATCCGTTCCAAGCGCTCTCGGGCGCATTAGGCTGTTTGTTTGCATTTAAGA
GATGAGTTGGAACGCGACTTACTGCTCGGAGAGAGAGAGAGTGCAAAGGCATGTCTGCTGTAGATGGAAGTCTCGCATTCTTTGGAATGGTTAGTTAAGCTTGCAGTGGTGACAAATCAAG
GCGGCCTGCTTTGGCGCTACGCATGCCCAATATAGTGATAGTTGCTCCATCTATGAATGGCAATAATGCACTTGGCATTAACCCCTTTAGAACATAAATTTCTTGAAATGCAACGTACCATCCTT
AATGAAACCGTTTCAGGAACTTTGTTAGCACCCGATATTTTGTTGATGCACCTATATTTATTCTGACGTCTTCGGATTTAGTACTCCCGGTTGCAGAAGTACATTTCCAAATGCGGAAATAGCAC
CACATTCCAGAGGACGTGTTTCGCCAATTGTTTCGCGAAATCTATTTATATCACTGCAAGTGGAGGACAGCCCTTTGAGGGATATGACAGAGCCTTACGTCGAGATGTATGCATAGTTTTATGT
CGTAAGACCGAAATCACACCTTCAATCGGGGTTTCTCATACTGTTTTCCATCCCCTGCTAGAATGTTTGCATCCCAGGACTCGAATACCAAACCCGTGTCCGTAATAGAGGATAGTTGGTCTGC
TTCCGAGTGATTTAAGTTCGATGCAGTTGTCCATGAAAGCCAACGAACCCTTTGATCGAAGACTTGCGAAGCTGTGGCAAGTTTTCAAAGATTGGGTGTTAGCCCATATATATATACTCTTCCA
TTTGAGAAACACATGCTTCCAAATATTGACAAATATGTTAGCTGTAGCAGTATTGCATTTTGTTCGTTTCTAGGCTCGACATCTGTACAGAGGAAGTTCTGGGGATCGCTCTAAATATTTATCCA
TGCTCGAATGCTCGGTAGCATTAAAGTACAAGTCTTATTATTGAAGGAGTCAATCATCGACGTCATATAAGCTACTATATACTTGCAGCTATTCGGCCAAGAACAAGTGTAAATAGCATTTGGC
TGTTGTCAGTAATGGGTAAATAGGGCACCGGAACAGCACGTGACACTAGCGCGGGCTCGTCGTCTTCGTTTTCTCCGTACGACGACATTCTCTCGCTGGCCGCTTGCGGACGATCAGTTCCAC
AACAATTGAACATCAACTTCCAATGCACTATCTCCTACCGTTTTCCTCGACTAGCTCTGCTTATCACTTTGAGTAGGCAGGCAAATTAACCGGATCCTCCGAGACCATGCAAGATCCAGGCCCAT
GTGGCCGGCCGATTATGCAGCACGGCTGTGGCTGCATTTATCAACACGCTCTTATAGGCTAGAAGACTTTGCCAAATATGGGCATTTTGATCCGTATGGATTCGGATCTGTGGATCGGCGGTT
GGTGGTCCTAGTCCGCTGTCTTCCATCGTAACGGTGGCATCTAGCAGCTCGAGGTTATCCACATTGCCGTAAATCGGAGGCGAGTTATGTATTGTTTTCCACCGAGCCTCCTGTCACTAACTTC
CTAATTCAGCATCTTGCCGTCTCTCCCGTACCCCAATCAGGAAACTTCACAATTCTAGAATCTGTAAATGCTCTTCAAGTGTCTTTCTCACGTTGCCGCTTGCAACGTCGCCTCACCAACGTCATC
GATGCCTTGACATAATTTGTAACCATGTTTGCGCATATGTTCTGGACATAAAGACCTAACTTGACTAAAGTTAAACCAAGGAATATCTTCCCTTTTTATACAACCTGGCGTCCTTTTGTTGCAAA
ACGTCTCCCAAAGTTCTACGTTGACAATCTTCGGCAACATTAACCACTGCCTATCAACAAGCCAGCTTTCAACCAAGCAAACTAAAAAGAAAAGACATTCATCCAACTGCTATAATCACTTTGGT
CCTAGCCAGCAATTAGAAAAAGAGCCGTTATGGGCTCGGCGACTTGGGCCTATGTCTTTGGTGACCAAACTAGCGAGTTTGACTCTGGGCTCCGACGCCTCATCCAGGCAAAAAGCAACAGC
CTTGTCACTTCTTTCCTAGAAAGATGCTTCTATGCTTTGCGCCACGAGGTCACCCAGCTTCCGCCATCTGACCGTCGGAAGTTCCCTCGCTTCACCAGCATTGTTAGTTTACTCTCGAGATATCGC
GAATCAGGCCCAAATCCTGCACTTGAGAGTGCTTTGACTACTATTTACCAACTGGCCTCTTTTATCAGGTGAGATACCCGTCATCATAGTAGTCTGATCTATATTCCTCTAATCGGTACCAGTCA
CTATGGAGATGGAGGCCAAGTCTACCCAGTAGGGTCAGAGAGCTATGTTATTGGACTGTGCACTGGTCTTCTTGCATCTGCAGCCATTGCCTCATCAAGGACAATTGGCGAGCTTATTCCAGT
TGCCGTCGAGGCAGTGGTCGTTGCCCTTAGGCTTGGGCTGTGTACCTTCAAAGTCCGTGAATATGTCGGCCAGTATGACGGCAAATCCCAAAGCTGGTCTGCTGTAATCTCAGGCATAAGCG
AGGACAAGGCTCTCGAAATTATTGAACAATTTTCTGTCAAGAACGTATGTTATGATCCTAGCCTGAATATTAGGATTTCATGCTAAACTTGCTCAGGGACTCTCCCCTTCATCTCGCCCTTACGT
TAGCGCCGTCAGCTACAACGGTCTGACCATCAGTGCACCACCTCTTGTACTTGAAGAGCTCATTAACTCTCAGTTGGAAAAGATCGCCAAGCCACTCAGAGTACCAGTTCACGCACCTTTCCAT
GCACCTCACCTCTATGGGCAGGATGACGTAGAATGGGCACTGAAAACCTCAACAAGCGGTACTTACAAGGGCTATACTTGCCGTATACCAATCCTCTCGAGCAATAGCGGTCGTGAAATTGTT
GCTGAAAACTTCGAGGGCTTGCTGAAAGTTGCAGTAGAAGAGATTCTCTTGAAACAGCTTTGCTGGGACAAAGTTCTCGATGGCTACACTTCGGTTATCAAATGTTCCTCGGTTTCTAAGTGT
ACCATTGTTCCCGTGGCGACCTCCGCTTCTTCTGGAATTATGTCTGCTTTGAAGCGCGCTGGAATTTCAGAAGTCGTGTTGGAAAACGCTCTTGTGAGCTCGTCAAATACTGAAAATCACAAGA
GCACCACTGGAAGGGCCGAACAGTCAAAAATTGCCATCATCGGTCTTTCGGGAAGATTTCCAGAATCAGACAGCCCAGAAAAGTTTTGGGATGTCCTATACCAAGGCCTTGATGTTCACAGA
GAAGTTCCCCCTGATCGGTGGAATTGGAAGACTCATGTCGATGTGACTGGGAAGAAGCGAAACACCAGTCAAGTTCCCTATGGCTGTTGGATCAACGAGCCTGGCCTCTTCGACCCGCGATT
CTTCAACATGTCGCCAAGAGAGGCACTTCAGGCAGACCCAGGCCAGCGACTGGCATTGGTGACCGCCTATGAAGCATTGGAAATGGCAGGTTTTGTACCCGACTCCACCCCTTCGACACAAC
GGGACCGCGTCGGCATCTTCTACGGCATGACTAGTGATGATTATCGTGAGATAAATAGCGGTCAGGATATTGACACCTATTTCATCCCCGGCGGAAATCGAGCTTTTACTCCTGGTCGAATTA
ACTACCATTTTAAGTTCAGTGGACCTAGTGTGAGCGTGGACACTGCTTGCTCATCCAGTTTGGCGGCCATCCATATGGCGTGCAATTCATTATGGAAAAATGACTGTGATACTGCCATTGCCGG
AGGCACGAACGTCCTCACAAACCCTGACAACCATGCTGGTCTCGACCGAGGTCACTTCCTTTCTCGTACTGGTAACTGCAATACTTTCGACGATGGAGCTGATGGCTACTGTCGTGCGGATGG
TGTAGGTACCGTTGTACTCAAGCGGCTCGAAGATGCCATCGCCGACAACGATCCCATTCAGGCTGTAATTGCAGGAGCTTATACCAATCACTCTGCCGAGGCAGTTTCAATGACTCGTCCCCAT
GCTGGAGCCCAGGCTTTCATCTTTGACAAACTTCTCAATGAAGCCAATGTCCACCCTCACGATGTTAGCTACATTGAGATGCACGGAACGGGCACACAAGCTGGAGATGCAGTAGAGATGAA
ATCAGTGCTTGACACCTTTGCTTCTGATTATGGTCGCAAGCCCGAACAAACACTTTATCTTGGATCAGCAAAGTCGAACATTGGCCACGGGGAATCAGCATCTGGTGTTTCTGCTTTGGTTAAA
GTATTGCTAATGATGCGCAAGAACACTATTCCTCCACACTGTGGTATCAAAACAAAGATTAACCACAACTTCCCCACGGACTTTAAGGAACGAAACGTCCACATCGCTTTCAAACCTACTCCAT
GGGTTAAGCCAGCGTCGGGATTGCGCAGATGCTTTCTAAACAACTTTTCGGCTGCTGGCGGAAATACTGCATTGCTTTTGGAAGATGGGCCAGCTGCTCAAGCCCAAGATGCTATCAAGGAT
CCACGCTCGACGCACCTAATTGCTGTTTCCGCCAGATGTCAGAATTCGCTCAAAACTAATATAAATAACCTGGTGGAGTACATCAGCAAGAACGCGCCAAACATGAGCGAGGATGGATTTCTG
GCCAAGCTGTCTTACACTACTACCGCTCGACGCATTCATCACCCTTTCCGGGTTATGGTGTCTGGTGCAAAGATTTCCGATATCCAAGATTCTCTCCGTGCAGCAGCTAGTAGAGAGACCATCA
GCCCTGCGCCAGCTGTCGCCCCAGGTATCGGTTTCGTTTTCACCGGACAGGGAGCTCAGTATGCCGGCATGAGTCAGCAATTTTTCGAGAGCTTCTCAGAATTCCGCTCTGATATTCTTCGTTT
CAATGGAATTGCCCAGAGTCAGGGATTCCCATCCTTCCTCTCCTTGATCGATGGGAGTTCGCCAATCGAAGAGTTGAGCCCAGTTATTACACAGCTTGGAACAACGTGCATGCAAATGGCATT
GGCAAATCTTTGGATGTCCTGGGAAGCTCGTCCCACTTTAGTAATGGGCCACAGTCTTGGCGAGTATGCTGCTCTATATGTCTCCGGAGCCTTGACGGCGGTTGACACCATTTACTTGTGTGGC
CGCCGTGCTCAGCTTTTAGAAGAGAAATGCAAAATTGGAACTCATGCCATGTTGGCAGTGCGGTCTGCCTTGGGTGAAGTGACTCCTTTCCTTAACCCAGGAGTACATGAAGTGGCTTGCATT
AATGCTCCAGGAGAGACTGTCATCAGTGCCACAAACCAGAATATTGATGCTCTTTCTGAGCAGCTTGTTGCTGCGGGAATCAAAGCAACGAAACTGAAAGTTCCTTTTGCATTCCACTCTGCTC
AGGTTGAGCCAATTCTAGAGAAGCTCACCGACGCTGCTAACGGGGTGAAGTTCCGCAAACCTTCTATTCCCCTCGTATCTCCTCTTCTGGGGGAAGTTATCAATGAGGCCAACTATGAGCAAC
TTGGTGGCAAGTACCTCAGTCGTCACTGTCGGGAGACAGTCAACTTTCTCGCAGCTATCGAGGCTTCTCGACATGCCAAGCTCACGACGGACAAGACCGTTTGGATTGAAATCGGTTCTCATA
CTATATGTTCTGGTATGATCAAATCTACTCTTGGCCCGCAGACAAACACCGTGGCTTCACTTCGCCGGAATGAAGATTCATGGAAGGTGCTATCTCAAAGCCTTACATCTGTGTACTTGGCTGG
AATTGAAGTTCAATGGAAAGAATATCACCAGGACTTCAAAAATGGCCACGAGGTTCTTCACTTGCCTTCCTACAGCTGGGACAACAAGAACTACTGGCTGTATTACAAGAATGACTTCTGTCT
AACCAAGGGCGACAACTCAGTACCTCAAAATTCAAACACTCCTCGTTCCAACCTCACAATTTCAGTGCAGAAGATTATTGAAAGTACTGGCGATGCGACAAAAGCAAGTGTCATCGCCGAAAC
TGACATTTCCGATCCCCTTTTGTGCCCTGTTATTCAGGGCCACAAAGTGAACGGGACACCTCTTTGCCCCTCATCTTTATATGCCGATATTGCCCAGACATTGGGAGAGTATCTCGTTGACAACT
ATAAGCCCGAGCTTAAGGGTATCGGATTAGATGTTGCCAACATGGCTGTCCCCAAACCTTTGATCTATAAGAATGCAGGGCCACAGCTTTTCCGTGCAGCTGCCACTGCAGATTGGGATGCGC
GCCAGGTGTCAATGCAGATATACTCTGTCACTCCGGAGGGAAAGAAGATCACAGATCACGCAAGTTGCGTTATCAAGCTATTTGACACTGAAGCGGCGCAAGAGGAATGGAAGCGCTACAC
ATACCTTATTCGCAGAAGTCTTGATCAGCTGTTCGAAAGCGCTTCCAAGGGCGACAGCAATAAGCTTGGTCCTGGAATGGTATACAAGCTATTCGGGGCTCTAGTCGAGTATGACAACAATTA

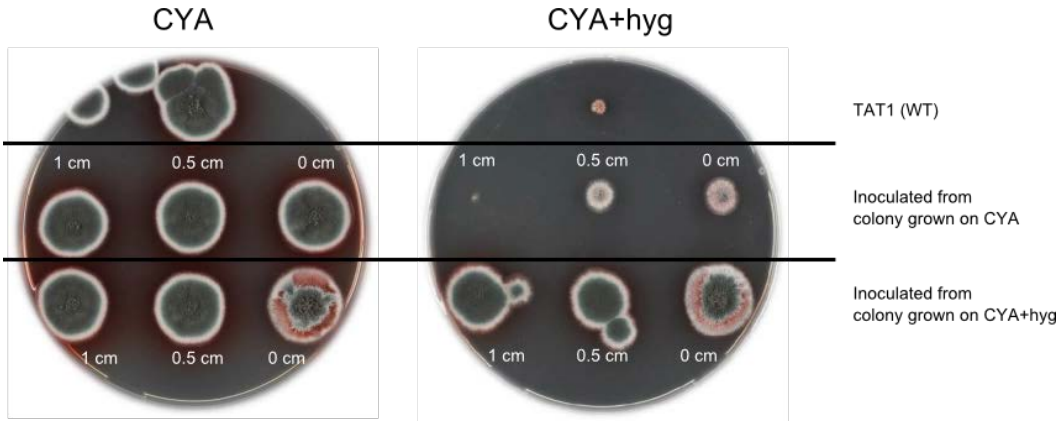
CAAATCAATGCGAGAGGTCATTTTGGACAGTGAACAATATGAAGCCACTGCTTTGGTAAAATTCCAAGCCGAGGCAGCCAACTTCCACCGAAACCCATTCTGGATTGACAGCATAGGGCATAT
TACTGGGTTTACTATGAACGCAAACGACGCAACTGACTCGGCCAATTATGTCTATGTTAACCATGGCTGGGATTCGATGCGCTGTGCAAAGAAGTTCTCTGCTGATGCATCGTACCGCACCTAC
GTGAAAATGCAACCCTGGCAAGGGACTATCTATTCCGGAGACGTCTATGTCTTCGAAGGAGACGAAATAATAGCTGTATACGGCGGTGTCAAGGTACGTACCATCTAGTTTGATAGTATAACT
AGAAGATGCTAACTGGCGATTATCTAGTTCCAAGGAGTACCGCGCAAGATCCTCAACACTGTCCTTCCACCACCAAGCGGTGCACAGGCAGCTTCAAGGCCATCAGCAAAGGCCCCTGTTCCT
GCTCCTGTCAATGTGTCAAAGGCGAAGCCCATTGCCAAGGCAGTCAAGCCAAAGACATCTGCTCCGAGTGTCCTCGTACAAGCTTTCAAGATACTGGCCGAGGAAGTTGGATTGTCCGAGGC
AGAACTTTCGGACGACCTTGTTTTCGCCGACTATGGCGTTGACTCCCTGCTTTCCTTGACTATCACCGGCAAATTCCGCGAAGAGCTCAGCATTGACCTTGAATCTAGCGTGTTCATGGATTGTC
CCACTGTCAAAGACTTCAAGCAAATGTTGTTGCAGACTGCAGCTCCCACCGATAGCAGTGAATCCTCTTCGGACTCCTCCGTATTCGATGACACTGGATCGGCTTCGGAGCCACCTACTCCGGG
AACTCCTGGAGACTTTTCTACACAAATGAAGACTACTCCATCCCAGGGAGAGGACAAAACGCTATCTGCTATCCGCCAAATTTTGCTTGAGGAGATCGGTATTTCCGCTGACGAACTGACTGA
GGAGGCCAATCTCAATGAGATGGGCATGGATTCATTGCTTTCTTTGACGGTACTTGGAAGACTACGTGAATCACTTGGTCTAGATCTACCATCGGAATTCTTTATTGAAAATCCAACCATGGGC
CATGTCGAAGTCGCGCTTGACTTGAAGCCAAAAGCTCCCGCTCCACAGCCGCCTGCAGCAGTGTTGCCTTCGGCGCCAGAATTGAAACCTGAGGTGTTGTCTTCTAAGCTCAGCCACCCTCCG
GCCACGTCTATTCTCCTGCAGGGTAGCCCTAAAACAGCTAGACACACGCTGTTTCTCTTCCCCGATGGTTCTGGCTCGGCTACATCATATGCGACGATTCCCAAGATCTCATCAGACGTCTGTGT
CTATGGGTTGAACTGTCCGTATATGAAGACTCCAGAGAAGCTGAACTGCGCCCTTCAGGATCTTACAGCCTCTTATGTGGCCGAGATGCGACGTCGTCAGCCCAATGGCCCGTATAACGTTGG
CGGATGGTCTGCAGGAGGTATCTGCGCTTACGATGCAGCTCGCCAGCTTGTTATCGAGCAGGGTGAAACTGTAGAGAAACTCTTCCTTCTAGACTCGCCATTCCCCATCGGCCTCGAAAAGCT
CCCTCCACGCCTGTACAAATTCTTCGACGACATGAAGATTTTCGGCGATGGCGAGAAGCCTCCACCGAAGTGGCTGCTACCGCATTTCCTCGCTTTTATAGACGCACTCGACATGTACAAAGCG
ATCCCATTCCCAAGCACTGACCCGCAATACGAAAGTCGTCTCCCCAAGACCTATGTCATCTGGGCCAAGGATGGTGTTTGTGGTAAACCAGGAGATCCACGTCCTGCACCACCTAAAGATGGC
AGTCGGGATCCCAGGGAAATGCTTTGGTTGCTTAACGACCGTACAGACTTTGGCCCCAATGGCTGGGACACGTTGGTTGGCAAAAACAAGGTCGCTGTCATTGAGTCCATATCGAATGCCAA
TCATTTCACAATGATGGATGGTGCTGGAGCACGGGGATTGTCTGATTTCATCAAGAATGCTTTTTAATTGTTTAGTTTTTGTTTTTCTCTTTCTCTCATTTTGGGTGTGTGGGAGTTATTGTTCTT
TTGGTTTTTTTGTTCTCCTTCAGGCTCAACATGTTGTCGTTCCTTACCACCAAAAGTGATACGCGTACTTGATACTATGACTACTACCTAGGTGTGTTTCCGTGGTGCTCTTATGGTATATATTTC
GGGTATTCATAATATGTTTTATAGTGATATGGGTATGGGATAGATAATAATTTCACTTAATAAAACAGTTGCCATGGTAATTTTCAGGAAACCTTCATGTAGGTATTCAATAAATCTTACCTAGG
TAACTACGAGGATAAACGTTATAGTCATCTGTGTCTCCATTTTCGACCATCTTTTGAAATTCTAGGTATTCTTGCATAAGATTAAGCCGAAAATCCATGCCAATTAGGGTATGTGATACTGTCAC
TGTCTCTATGCATAAGAATAGCTTCTACCGAAAAGTTGAACATATGTATAACAGGAAGGACATACGAGGTATAAGAGATGAAAATCGAAAACAATATTTCAAATGCTCAAACACTCAATTGAG
GAACCATGATTACTGAAACAAACACCTTCACGCCCTCGTATTCTTTGCACGGGGAAGAGGAAACAAGAAAGGAAAAGGACAAAAGCAATACAAATGACACGCTAACATGGATCCAACCAAG
AATCATTCTTGATATCATCCCAAACATCGTAGCCATCCAACAGCACCAAGCCCATGCCAGACATCAAGTGGGTTTCCATGTGGCAATGGAGCAAAAATGGACCAGGGTTCTGCACGAAGTACC
TTATTGCTAACCACGACTCTCCTCTAGGCGACGTAACGACAGTGTCTCTATACACAGGCTTTTCAAAGAAATCTTGTGGACTTTCAACGATTCCCTGTTCGATGGAAGACCAGTTAAAGGTCCC
CTGCCCGGCACCCATCACGAACATCTTGTTTGAATGTTTGTGAATTGGATGAGGAGGCTGAACGGGGGTGACCGATGGATCTCCTAGGCGTAGTTGGTAGACGATGTCGACCCAGGTGTTAT
TCCTGGTAGTGATTTTGAGAGCGTTGTCGAGTCCTGGAGATTCCTTGTTGAACAGGACGGGTTCCATGCTATCCAGATCGGCTGGAAGGAACGTGGTGCCGCCAAGAGACCATTCCCAGGAA
GAGTTGATGCGCGATAAATTCAAGATATGGAAATCGTCTGCCTGAGGCGCAGGGAGTATTGCTGGGAAGGGAGGTAGGTTGGTGACATTCAACGCAGTAACGTCTGGTGAGGTAGGTTGG
CCGCCGTAATTGAGGTAGGGCACGGATGGCCGTGTGCTATTCTCTCCACCAGCATACTGCAAAGTCGCATAGCCCGCAATTATCTGGTTACCGCCGGTGTCCGTGATGGTTATATTGTAATTG
GCTGGTTCCTTGTCGAGTTTAATGAGCGCAGAGTAACGCTCGCCATTGTAGAACTCGAATGTATCAACGTATTGAGGCTCTATGTATCGGCCATCGACCTCATAAATCCACATTGGATGTTCAT
CGACGGAGACAACCAAGGCTTTCTCGCTGGCTGCGCCGATGAAATTGAGACTAGCCCATCCATGAGCTGGGTCAACTTCTATGACCGCCTTGTACCCATTGCTAGGGATGCAGCCATCGTTAA
GATAAGCAGGAAGCTTTTCTGGGTGGTGTTCCCAGAGACCCTGAGTGCCGGTTACAAAAGGGAGGCATCCCTTGTCCGTTACGTTTGCCTGAAGAACATTCTGAACGGACTCTGGCATCAGA
GAGATTAATTCCGACTGTGGAGGGCAATAGACGGCGCCCTTGCCTTGGATCAACACACTGTCCATACAGCTGCAGGTAACATTGTTAGTATCGATGGACTGGAACAATCGCACAGAGAATGC
CACTTGGAGCGCGAGGGGAATTAAAGCAATAAAGGTAGAATGCAAGGGATGGAGATTCCGTACAAGATGTCATAGCCTGTTTGCTTCAGTGCATCCATATAGTCAACAGAGGTGTATCGACT
CCAATCAGAAACGACGATAAGTTCGGCATCATGCTCCGCTTGCTTTATTGCTCTAAGTTCGTTCTCGTCTGTTGTAATCATCGAAAAGGGATCTTCTCTTTCTGGCGATGGCCTGCATTGCTTCA
GTTAGCTAGAACAAGCAAAATTCTCTAGTGGCACCGCCAAATTTTCTTCTGGCATAACAGGCTTTTTTGAAACTAGGGATGCGGTACGAGAATAGAAATGGGACCGAGGATGGGTGTTCCAA
ACACTTTTCTGATATGAACTGAAATGAACTAAGATACTCACCGGACACGGACTGCCCCATATAGGCCGTCTTGCAGGGTTGCCATATCGTGCGAATGATACCAGTAAGTCCCATAGATATTATT
GGCATGCCACTGGTAGGTGAACGACTCTTCGGGTTTAATGGCCCACTGGGTCAAGCCAGGTACACCATCTGCCCACGCCGTATTCTTCTGCCTATATGAGGAATAATGAGTATTGGAAGGTCT
TCAGCAAGTGAAGGGAAAAAAAGGGCGTCAATCACATACTCGATTCCGTGGAAATGCACTGTCGTATTGAACGGTAAATTGTTTTTCACCTTCACTGTAACATCGTCGCCCTCATCCACAATCA
ATGGAGGTCCCGGAAACTGTCCATTGATAAAGACCATATCTCTCTCCTGTCCATTGGGCGCTCCGAGCCCCCATGTAAGTTCGAGATCGAAATGGACATGGGCACTTTGTGCCAGATTCGCCA
TAGCGGCCAAGGGCAACATCAATCGGCTGGAAAGAGCCATGATGTGTCTTTGGGCTGGTTACTAAACTGATCAAATCGACAAAGAAGGTGACTATCAAGCAACGAGTCAAGGGAAAAAAAG
TCACTGGCAGGAACTCGAAGTAAAATGGAATGAATTCCAAACACTCGACTTAAAAAGAGTTTCTGCATCGAAATGCAAGTTTGATATACTACGCCTACGGCCAATCCCTTGC

S1 Fig. The stability of AMA1 plasmids in *Talaromyces atroroseus*. To follow progressional loss of the AMA1 plasmid as the fungus grows in the presence and absence of hygromycin selection, we used the following experimental setup presented in (A). Step 1, *T. atroroseus* was transformed with a CRISPR plasmid (pFC332) containing the hygromycin resistance gene, but no sgRNA gene insert, and plated on solid CYA medium containing hygromycin. Step 2, Spores from the resulting transformation plate (CYA+hyg) were transferred via a single-point inoculation to a CYA plate as well as a CYA+hyg plate, and allowed to grow for four days at 30°C, forming colonies with diameters of approximately two cm. Step 3, from these colonies, spores were collected from an area of approximately a mm^2 from positions with increasing distance from the centers of the two colonies and transferred to solid CYA medium with and without hygromycin. Specifically spores were recovered from the following distances from the center of the colony: 0, 0.5 and 1.0 cm. (B) Solid CYA medium with and without hygromycin from step 3. Reduced growth on the CYA+hyg plate as compared to growth on the CYA plate indicates loss of the AMA1 plasmid.

A



B



S1 Table. List of primers

UA08_00425-up-FU	ML571	GGGTTTAAUTCACTACGTTCCGATTTTCGC	UA08_00425 deletion in T. atroroseus
UA08_00425-up-RU	ML572	GGACTTAAUAACGGCTCTTTTTCTAATTGCTG	UA08_00425 deletion in T. atroroseus
UA08_00425-dw-FU	ML573	GGCATTAAUCTCATTTTGGGTGTGTGGGAG	UA08_00425 deletion in T. atroroseus
UA08_00425-dw-RU	ML574	GGTCTTAAUACTTATCGTCGTTTCTGATTGGAG	UA08_00425 deletion in T. atroroseus
gRNA-Tatro00425-1FU	ML559	AGTAAGCUCGTCAACTTCCGACGGTCAGATGGGTTTTAGAGCTAGAAATAGCAAGTTAAA	UA08_00425 deletion in T. atroroseus (protospacer 3)
gRNA-Tatro00425-1RU	ML560	AGCTTACUCGTTTCGTCCTCACGGACTCATCAGAACTTCCGGTGATGTCTGCTCAAGCG	UA08_00425 deletion in T. atroroseus (protospacer 3)
gRNA-Tatro00425-2FU	ML561	AGTAAGCUCGTCAGCTCTTCAAGTACAAGAGGGTTTTAGAGCTAGAAATAGCAAGTTAAA	UA08_00425 deletion in T. atroroseus (protospacer 2)
gRNA-Tatro00425-2RU	ML562	AGCTTACUCGTTTCGTCCTCACGGACTCATCAGAGCTCTCGGTGATGTCTGCTCAAGCG	UA08_00425 deletion in T. atroroseus (protospacer 2)
gRNA-Tatro00425-3FU	ML563	AGTAAGCUCGTCGAGAAGTTCCCCCTGATCGGGTTTTAGAGCTAGAAATAGCAAGTTAAA	UA08_00425 deletion in T. atroroseus (protospacer 1)
gRNA-Tatro00425-3RU	ML564	AGCTTACUCGTTTCGTCCTCACGGACTCATCAGGAGAAGCGGTGATGTCTGCTCAAGCG	UA08_00425 deletion in T. atroroseus (protospacer 1)
talA-up-FU	ML457	GGGTTTAAUCTCACATTACTTTCATCTGGTCTCG	UA08_04451 deletion in T. atroroseus
talA-up-RU	ML458	GGACTTAAUGACAGAATATATAAAATCAGAAAAGTCAAATAC	UA08_04451 deletion in T. atroroseus
talA-dw-FU	ML459	GGCATTAAUTCGTAGATCGGGGCAGAGC	UA08_04451 deletion in T. atroroseus
talA-dw-RU	ML460	GGTCTTAAUTGCATAGTTTGTCATTTTATCCCA	UA08_04451 deletion in T. atroroseus
gRNA-talA-FU	ML461	AGTAAGCUCGTCAACTAAGCTAGAAGAACAGGGTTTTAGAGCTAGAAATAGCAAGTTAAA	UA08_04451 deletion in T. atroroseus
gRNA-talA-RU	ML462	AGCTTACUCGTTTCGTCCTCACGGACTCATCAGAACTAACGGTGATGTCTGCTCAAGCG	UA08_04451 deletion in T. atroroseus
UA08_00425-upchk-F	ML581	GACCAGTATCTTCGTATCTTTCGC	UA08_00425 deletion in T. atroroseus - check primer
UA08_00425-hkchk-F	ML582	GATGCCTTGACATAATTTGTAACCA	UA08_00425 deletion in T. atroroseus - check primer
UA08_00425-hkchk-R	ML583	CCTCCATCTCCATAGTGACTGGTA	UA08_00425 deletion in T. atroroseus - check primer
gpdA-p-int-rv	CSN105	TTGGACGCCCTACAGATGC	Deletion up check primer (hph promoter)
talA-upchk-F	ML472	ACCAACGAGCATAGTATCTCCGA	UA08_04451 deletion in T. atroroseus - check primer
talA-hkchk-F	ML515	TGACCCTCTGGTGGTAGTTGG	UA08_04451 deletion in T. atroroseus - check primer
talA-hkchk-R	ML516	GTCACCGCACATCAAGCCTAC	UA08_04451 deletion in T. atroroseus - check primer
TATR-albA-Dw-Chk-R	B409	CTATTCTCGTACCGCATCCC	UA08_00425 deletion in T. atroroseus - check primer
TATR-talA-Dw-Chk-R	B410	CCATTCCATGCAACACCTACGATC	UA08_04451 deletion in T. atroroseus - check primer

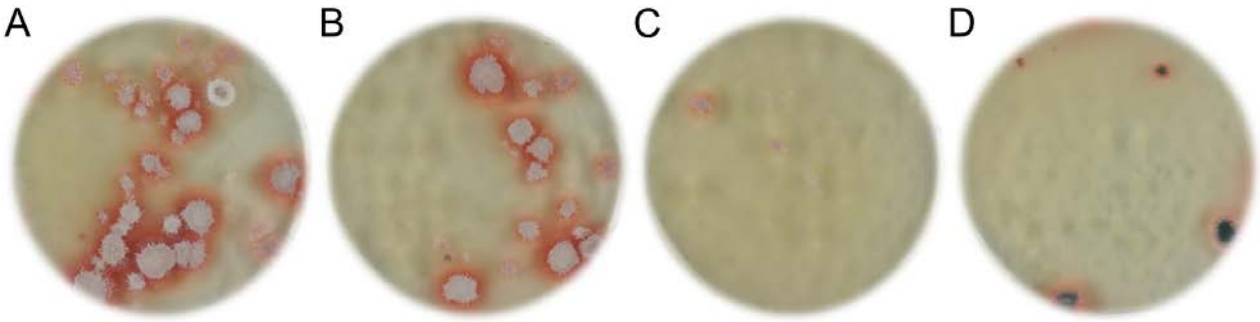
S2 Appendix. DNA sequence of *talA* including 3 kb up- and downstream sequences. The coding sequence is highlighted in yellow, while predicted introns are marked in green.

CCAGCGTTCGAAATATAGTCAATCCCACTGTGCCAACGACGCACCATCCAATGGCAGCAGCTTCGTCCCAGGCCATATCATCCGGTACTCTCCACAGTTGGTCTGCCCACACTGCGATGTATTC
TGCAAACGCGCCAGACTGAGGATCTAGACGGTTCGACGCATGCACGGCTCCTGCTACCCGGTCGCCTATCTTAAGAAAAGACCTGTCTACGCTGTCGCCCAAGGCAATGACGGTGCCTGCAT
AGTCCGAACCAACACCTGCACCAGGACATGGGAAATTGGTGGGCATTTTCCAGTCACATGGGTTTAGAGCCACTGCTGCTGTCTTGACGAGCACTTGGCCTGGCAGAACGCTAGGAAGACCT
CGATTGTGACATAGGGTAAGTGAGCCTGGTTCTTCAGCTTGTATGACTGCTCGCTGGATTGTTGGAACATCTAATTTAGCCGCCATGATTGATCGGGGGTTGGAAGGTAATAAAAGCGTATTG
TGCACGGCGGTAGGGATACTTCAATCAATGAACTCCCACTGAATATCTTTTGAAGCGTGACCGACTAGTTTATCAACGTCAAACTGACTTGGCACCTTTCTGTATATCTATCCTATCTCTACGGT
CCCGAATATACTTGCGAACATACACTAAAGAGGAGGCTGTATTACTTTCTTATTTGTCTTCATGTTTTTTTTTTTTTTTTTTTTTTTTTTTTCAACTCTTTTGATCCACTTCCCAACCAACGAGCATA
GTATCTCCGAATATCAGCTGCAAGAGGCTTGCTCCGAATATGAAATTCGGACCAGGGAAACGCTCCGAACTGAATATTGTCAGAACTGATTTCCTATCTGGACAATAATTACAGGGCAATTCT
AATTTCTGATGAGGCAGCCTATTGCAGAGGTAACCAAGTTGCTGTTTGATAACATTTTTCTTTCTTTTCACTTTCTGTTTCTTTCTTTCCGGTCATTTGTCTCACAAATTGTCCTTGCAAAGTCCTG
TCGATAGTATGTATTACTTTCCGGTCCTTCTCACATTACTTTCATCTGGTCTCGCGGAGCAATGGCGGCAACAGCCCCTAGGAGATCTATTTTCAGGGCACTGGGAGAGTCCGTTGGGGGGATT
TGGGCTCTCTCTATCAGAATGCAAAACCAGCCATACAACAGCATATCAGCAAGGGTATGGGCCGGTAGAATTCTCTCTCTCAACGACAGAGCATCTAGAGGTACCAAAACTATCGTCCATCAA
CGCCACAGCATGGGAACAATGGGAATTTGACGGTGTGTCCGAGTCTGGCAACGCCGGATTGATTGTGGGGTTCTCGCGTGATGCCTCCTACGCCTTCTTTGGTCAAGGCAATCTCCGCGTTGA
GTTTTTCATGATTTTGGAGGATGGAACTGTCATTCAAGAACTTGACTACCTACAAGAATCGACCATTATTGTCTGTGATGACACCGTAACGGGCATCTGGAATAGTACCGGCCGCAGTTATTCC
TTCCAGATCTCCAAGGACATGTCTCAGGCGAAAGTATGGTGGGATACTCCTCAGGCACAAGGAAGCCTGTCTCTACTGTCCCAATCGCCCCCGCATCTTGCAGATGGGTCGCTGTGGCCTGCT
GCTGACAAGAATGCAACAACTGAACTCGGGCCTGGGTTACACTTCAATCAACCCATCGCTGGCGGCCGTATAGAGGCTGATGTCACGATAAGAGAGCAGCGGATCCTTGTCTCGGGAACCG
GTGGACATGGCCGCCTCTGGGCCAAAGATGGTTGGTTCCAGATATGCAGTGGATGGTTCATTGTTCGTGCCTCTGCAGGGCCATATGTCCTTTCCTACTGGGAGATCGTATCCCGAATCAATC
CAGGTGTCTCATACTATTCTGCACAATTGTTCAAAGAAGGTGAATTGCTAGTCGCAACGCAGATTGGAGAGCGATCTGATGCTCAGGACCACGTTCTCTATACACAGGAGCTCGGAGGCAAT
GTGACCGGACTCCTCAGAGACCGCTCCACCGGCCATGTGGTCGAGTTTGTATCTCCAGTGCAAAGAAAGAGGTGGCGCTTCCTTGTTCGAAATAAAAGAAAGAAGTTCGAGATGGGCATGG
GTGGCGGCTCTGGCCTCACAGGATTCACAAATATTGTGACGGGAGGAGAAGTCTGGGAGAAACCATACGAAGGACGCGGATTCTCAGAGCAAGTGACCTTGCCGGACAAAATCAAACAATG
GCAAATTTGGCTTGTCTACGGTATTGGCTTCCTCAACAAGGGGAAAACGTCTTTGATCAAGATGATTGGATATATAATTTGATGCGGGTTTGGTTCAATTTTGGATCGTAGAACGCGAATGGT
AACTGTAGTTAAGTCGAATGATATGCTGGTTACGTAAATGAATCATGCTACTCGGTAGAATGATTCTGTGGGTTGATGAAGCTTTATATCTCTAAGCTATTCTCATATGGTGTTGTCACAATGG
TTTCAGGATGACCACGCTTTTTGACACTTAAAATCAACCAGCTCTGATACTTCTGTACTTTATGATACCCAAGGTTATCGCCATCTACCCCAAATGACTGGAATCCTGTTACCCCTGGCAAAGGA
AACAGACAAAAGAGAAATTTTAGCCTATATTCAATCATCCATCTCAGCTCTATGGTACGACATCGGAGATTCTCCATCATTCGGAGCCCCTTCATGATTCGGAGAGTTTGAAGAATGGCAGAAT
ATCGGTCCGAAACTGCGGGATGATTAGTATTTGTATGTAATAGAAATATCGGAGAAGGTTGGTGTTGACCCTCTGGTGGTAGTTGGATCTTAGAATAGTACATTATGGTCAGATTAGATGCTA
TAAACCTGGACCTTCTTAATTTCAACTCAAGTGATCTTTATAGTGGCCTTACTCCTTAAATTGATCTACGGTTATACCTAGGTGTATTTGACTTTTCTGATTTTATATATTCTGTCATTAAATAATCT
ATAAAAAAAAAAAGGCCTTCCAAGAGCATCCAACATGGGCACCATCCATACGCCTATTAAAGAGGCCAATGAGCCAATTGCCGTTGTTGGTAGTGCCTGTCGATTCCCGGGATCATTGAATAC
TCCATCCATGCTTTGGGATTTTCTGTGCAAACCTCATGATCTCCTTACCAAAATTTCCAACGAAAGATTCAACCCAGACGCATTCTACCATCCAGATGGAATGCATCACGGTACCTCTAACGTGA
AGGAGTCATATCTCCTTCGTGAAGATCACCGCGCCTTTGACGCCGGCTTCTTCAACATTAAGCCTGTCGAAGCTCATTCAATTGACCCACAACAAAGGTTGCTGATGGAGACAGTCTATGAGTC
TCTCGAGGCTGCCGGAATGTCTATAGAATCACTAGCTGGCTCACAGACGGGAGTATACGTAGGCTTGATGTGCGGTGACTTCAGTGAACATCTCCAACGTGACCCTCAGTCATTGCCTACATA
TATGCCCACAGGAACGGCTCGCAGTATCATATCCAACAGGGTGTCTTACTTCTTCGACTGGCATGGACCATCAATGACTATTGATACCGCCTGTTCTTCTAGCTTAGTTGCTGTCCACCAGGCA
GTGCAGCAGCTGCGAAGCGGCGAGTCAGATGTTGTTATTGCAGCTGGTGCCAACCTTATTCTTGGACCTGAGCTGTACATTGGGGAGGCCAAATTGAAAATGTTATCTGCCGATAGCCGGTC
GCGAATGTGGGACGTCGATGCTGATGGATATGCACGAGGGGAAGGAATTGCCTCGGTAGTTCTGAAAAGGCTCAGTGCTGCTCTTCAGGATGGTGATCACATCGAGTGTGTGATCCGCGAA
ATTGGAGTCAATCAAGATGGCAGAACAAAAGGGATCACAATGCCAAATGAATTGGCTCAAGCTGACCTGATAGAAAGAACTTACTTGAAGGCAGGCTTGAATCCTAAAGACAAGAACCAGC
GCTGTCAGTTTTTTGAAGCCCACGGGACTGGCACTTCAGCAGGCGATGCTCGTGTAAGTCTTGAGTTCCCTCTCTCCCCCCGCCCATTTTACAGTTCCATATATCAAAGCATGATAGCTGAGAC
AATAATAGGAGGCAGAGGGTATTAGTAGGGCATTTTTTGGCCATGAATTTGGCAATTCTGACGTAGAAAATGTGCTGTATGTGGGCTCGATCAAGACAGTGATAGGTCATACCGAAGGCACT
GCAGGCTTAGCAGGGCTCTTGAAAGCCTCCCTCGCCGTTCAGCATGGAATCATCCCACCAAATATGCTCTTTAATTCTCTCAACCCAACAGTCAGCCAGTTTTATAATAACCTAGAAATTGTGAC
AACCGCTAAACAATGGCCCAAAGTCGAAGGACCTCGCCGAGCCAGTGTCAATAGTTTCGGCTTTGGGGGCACAAATTGTCACGTGATCGTTGAGAGTTTTGAAGATTCAAGCAGCGAACACG
AGCAGGTGAAAGAAGTGCAAGTGCCTTTCACTCCTTTTGTGTTTTCTGCTGCATCTGAAAAAGCGCTAGAGGCAAATTTAATGGCCTATTCATCTCACCTTCGATTGAATCCCGATCTCAATCTC
GGTGACTTTGCTTTCACGCTACACTCAAGAAGATCAGCGCTCGGGGTACGAGCAGCCTTTGCAGCTAGGTCTTCTGATAGCCTCTGTGCGGCTATAGATGAACGCATTCAGCTCCATAAGACG
GACACGAGCACGCCACTCGGTGTCAGACCTAACACCGCTGCTCCATCTATCTTGGTTGTCTTCACTGGACAAGGAGCGCAATGGCCAGAAATGGGACGTCAACTGATATTATCATCGGCCTTT
GTTCGTAAATCCCTAGAAGAACTGGATCAAACGCTTCAAGATTTGCCAGATGCAGATAGACCTTCATGGTCTCTGGTGGAGCAGCTACTTGCTGATCGCTCTCACTCGCGCGTTGGAGAGGCT
GCCATCTCACAACCCTTATGCACTGCTCTTCAAATTGTTCTAGTAGAACTCTTAAGAGCTGCTGGGATTAGTTTCAAGGCTGTTGTAGGTCATTCATCAGGTGAGATTGCCGCGGCATATGCAG
CTGCGTTCATTTCTCGCGAGGATGCGATCAAGATTGCATTCTATCGAGGTCTAGTCGCGAAATACGCGGGACAAGATCGACCAGGCGCGATGATGGCAGTGGGGACATCATTTGACGACGCC
AATGACCTCTGCGCTTTGGACACCTTTGAAGGTCGTCTTTCTGTTGCGGCGTGCAATTCCTCATCCAGTGTGACTTTGTCTGGAGATGCGGATGCCATTGAGGAGGCTAGGTTCATTCTTGCTG
AAGAAAAGAAATTTGTCAGAACCATAAAAGTGGACAAGGCCTACCACTCTCATCATATGCTTCCTTGTGCGGAGCCCTACGTTAGTGCGCTTCTTGAGTGTAATATTGAGCCGAAACAGCCAA
AGACTGGCAGTTGTCACTGGTACTCGAGTACTTACCAGGGAATTAAGATGGAGGCTCAAGACGAGCTCAAAGGAGTGTACTGGAAAGATAACATGGTGAACCCCGTGCTCTTTTCACAGGCT
ATAGAGGCCGCAGCATCCAATGAGGATCCTTTCAATCTTGTTATTGAGGTTGGTCCACATCACGCCTTAAAAGGACCTGTTCTTCAAACATTGCAAGATACTTATGGCCAATCTATCCTTTATAC
GGGCCTCCTCAGTCGGGATAAACATGATGTGGAGAGCTATGCCAATGCACTAGGTTATTTTTGGAGCCACTTCAGTCCGACTGGTGTCGAACTTGGCCAGTACGACGATTGTCTCTCTGGAGG
TCAAAAGAGACAAATTATACCCTCATTGCCGTCCTATCAATGGGATCATGACCGGACCTTCTGGCACGAAACCCGCTTATCAAGAGCTCATCGTAACCGCAAGGATCCATACCATCCTTTACTG
GGTTCGAGAACACTAGAAGGGACAGAATCAGAAATGCGCTGGAGAAATATACTTCGGCCTAGCGAACTTCCTTGGGTTCATGGACACCAACTACAAGGCCAGATGGTGTTTCCAGCGACAG
GATATATTGCCACAGCTATAGAAGCTGCAAAATACATGGTCAAAAACATCCCTATATGTCTAATTGAAGTGTGTGACTTTGTTATTGGAAAGCCCTTAACGTTTGATGACGACCAGTCCGGCGT
TGAGATGTTGTTCACATTGTCCGATATTTCTAAGAACAGTTCGGGGTTGTATTCAGCCTCATTCAGGTACCACGCATGGACAAATCAGGAGTCAGATACATTGTCATTGTTAGCCTCTGGCTGC
ATTGTGGTAACTATAGATGAGACATCATCTTCGAATGATGTGCTTCCACCACGAAGTCCTGAACAACCGAATATGGCATCTGTGAGGGAGGATCAATTCTACGCTTCTCTTGACGACCTTGGTT
ATGGCTATTCTGGGGATTTCAGATGTCTTTCATCAATGAAGAGGAAGCTTAACTACGGTTCCGCCTACGTATCTGTGCCACAGCAGGAAGCCGTTGACGCTGTTCTTGTACATCCCGCCTTCCT
AGATGTCGGCTTCCAGTCCATATTCCTGGCCTACGGGTGGCCCAACGATGGAAGCCTTGATCAACTACATGTCCCTAGAAGCATCTCAAGGATCCGGATCAACTTTGCTCTCTGCCAGAAGGA
CTTAGTTCCTGGGACACAGCTACAATTGGACTCTCACCTGACAGATAATCCGTTGACCGCGGGCGCTATACGTGGGGACGTTGATATATTTGGAGCTGATGGACAATCCACCCTTATTCAGGT
TCAGGGTATACAAGTGGTTGGACTTTCAGATGGAAGCCCTCAGTTAGACCGGCAATTGTACTCAGAACACATTTGGGGCCTGGCAATCCCTGACGCTCAATGCGCAGCAGATAATCGCGCAA
CACCGGAGGATTACGAGCTTGCATGGAGTCTGGAACGTGTATCAATCTTCTACATGAAGAAATTACAGGCTGATATCAAGCCAGAGGAAAGAATAAACCTGGAATGGTATCACCATGCACTT
TTCGATTTCGTTTCTCATATTCTTTCCAAGACAGAGAAAGGCAAACAACGTTTTGCAAAGAGAGAATGGCTTGACGATACTTGGGAACAAATATTAGAGATCATGACACGGTAAGTACCACTC
CATATTTCTCCATGACCTTATAGGTACCTATGCATGTGCCAATTTCTCATACCTTGCACTAGATACCCTGATAGCATAGAAATGAAACTTACTCGTACGGTTGGCGAGAATCTCGCAGCTGCTGT
TCGAGGCGAGACAAACATTTTACAGCATTTACTCGCCGACGGTCTCTTAAACAAGTATTATACGGACGCGATGGGTCTCAAAGAAACAACAGAGTTCCTAGCCAAGTCTGTTGCCCAGGTTGT
GCATCGCTATCCACATATGAAGATACTTGAGATTGGTAAGAAAATATATATATATCAAATAATTCTATTTGTTAATACCTTTGTGTTCTAACTTTGCATCTTAGGCGCCGGTACTGGTGGAGCCA

CGAAATCCATCATGAGAGACATCGGCCGGTCCTTCTCTTCTTATACATACACGGACATTTCAACTGGTTTCTTTGAAACTGCCAAAGGAATCTTCGCTCAGAATCTGGACAGAATGATTTTCAAA
GCTTTGGATATAGAGAAGGATATACTTGAACAGGGCTATGCAGAGCATTCTTATGACCTTATCATAGGATCCTTGGTTCTACATGCTACTTCTAGCCTGCGGAAAACGATGGAAAATACACGT
CGACTGCTGAAGCCAGGAGGCTATCTCATAATACTTGAGATAATCAGCAATGATGTGATCAGAACAGGATTTGCTATGAGCGGCCTGCCTGGTTGGTGGTTGGGTAGAAATGATGGCCGTCG
TTTCTCTCCGTGTGTATCTTCCGCTGAGTGGCATAACCTGCTTGTAGAGACCGGGTTTTCAGGAATAGACTCACTCAGCCCTGAAACTGACATCTTACCACGTCCTTTGTCTGTGATCATATCTC
AAGCTGTGGATGATCAGGTCAATCTTCTCAGGGAACCTCTTCAATATCCAGGCCAGACACTAGGGAACACACAAGATTGGGATCTTGTAATCATTGGTGGAGAGACGCTCCGAACGATAAAA
CTAATAGATGAAGTAATCCGCCTGCTTCAGCCCTGGTACTTGCCAGTTACACGAATCCGCTCTCTGCACGATACCGAACGCTTGTCCAATATATCAGCCAAGAGTGTCGTGCTAAGTGTTACAG
AGCTCGACCAACCCATATTCCTGGATTTCAGCCCTGAGACTATGGAAGGTCTCAAAAGGCTACTTGGCTATCAGCGAACCATTCTCTGGGTCACCCAAGGTTGCAGAGTTGATGAACCTTACAT
GAATATGACCGTTGGGTTTTGCCGAACCCTGGCCCTTGAAGCTCCAGATACCCAGCTGCAGCTCCTAGATCTAGACATATCACGTAGGCCTGATGCTCGACTGCTTGCCGAGTCACTCTTGAG
ATTAAACTTCACACAGGTTCCCGAGGTCTTGTGGTCCGTTGAGCAGGAGATAGTCCAAGAGGCTGGGAAACTGTTAATTCCGCGTCTTATTTTGAATCAAGAGCAAAATAATCGCTACAATGC
CGCAAGAAGGACCATATTCGACACCAAGGATGTGCGCCAAACTCCTGTTATTCTACAACCCACACCCGTCGGTTATAGTCTTGTCAAGCACGAGCTGGATTTGTGCCCGAATGACGTTCAGGT
CGCAGTTAGTCACTCAACGTTGGTACCGATTGCAAATTCAATGTACGGTATTGCTGGGATAAATATATCCACTGGCAAGAGCGTTTTGGGCTTCTCCACTATCAATGGTTCCCAGCTTGCTATA
AATACTGACAGGCTTATTGAATACAATTCATCCAAAATAGGGGAGAGCAGTCGCCTGATCTCTCTTCTGTTGGTTGAGATGGAGGTGGACAACATCCTTTCAGTTTGTACATCCGATGAGCAT
ATTCTGCTCCATCAAGCATCTGCAGAATTAGCGGAAAGACTTGGAGAGCGGGCCCATGAGAAAGGGCTCAGAATATCCTTTACTACGTCTCATGGAAAGCCTCTGGACAAGACCTGGATTAC
CGTCCATCCATGCTCTCCACTACGTACTATTAAGGCAACTGTGCCTCTAGAAGTTTCGATTTTCATTGATTATTCAACCAACGACGATGGCGTTGGTGAGTTAATATCGTCCTGCCTTCCTGATA
CGTGTTGGCGAATTACATCTACGGACGACAGATTTCGTACCAGAATATCGGGGCAAAAACTTTTTGAGAAGCTCAAAGATTTCATTAAACGCGTAGGGACGGTAGACATTGATTCTTCCAGCC
ACTCGTCCCCGGTTCAAATGAGTCCAATGGATATTGCAGGGAACGGTAAAACTGCAAGCTCCACGATTGTCAATTGGAGTGGCACCACTGAGGTCTCAGTGAAGTTGTCATCTATCGATAGCC
AAATCAACTTTGCACCTGACAAGACATATGTGTTATTTGGGTTGACAAGCGATCTGGGACAGTCACTGTGTGACTGGATGGCTGCTCATGGAGCCCGGAATATTGTCCTAACTAGTAGAAACC
CGAAGATCGATCAAAGGTGGTTGGACCAAATGAAACGCACGGGTGTGAGGTTGCAAGTCTTCTGCAAGTAAGTGAGCCTCTCTTTCCTGGGGAAAAAAAAATGTACGCTGTCAAGAAGCTA
ACTTACGTTTTACAGTGATATAACCGATAAAGTGGCAGTCCAAGCAATTGTTTCAGAAATACGTGCCACATTCCCACCAATTGGGGGTATTGCCCAAGGAGCAATGGTATTGGAAGACGCTTC
ATTCTTTGACATGTCATTTGAAACTATGGACAAGGTCTTGAAACCGAAAGTACTAGGTAGTATTCACCTTGATGAGCTATTCCAAAACGATGACCTTGACTTCTTCATCTTCTTCATCACTAG
CATCGGTAAGCGGAAATCGCGGTCAATCCAACTACAGTGCTGCGAACATGTTCATGACCGCCTTGGCTTTCAACGTCGACGAAAGGGGCTCGCCGCGTCTGTTCTCCATATCGGTGCCATTA
TGGGTGTTGGCTACGTTATGAGAGAAGTGAGCGAAACCGTCTTTCCTGCGATCCGTCGTGCCGGGTTCACGTGGATGCCTGAGCGTGGCTTCCACCAATGTTTCGCTGAAGCTATTGTTTCTG
GACGGCCACAATCTGGTCGCAGTCCGGAGATTGTCACTGGACTCCGACTAATCAATGCTAACGAGGAAGAACCTGCTCCTTGGATGAATATCCCCAGATTCCAACATTGCATTGACCGCAGTG
GTACGGCGGGACTTAGGGAGGGTCATGGAAATGCAGTCGTAGCAGTTAAGACGCGGCTGCTAGTGGCATCAACTCGGGAAGAAGCCCTGGAGATGATTCAATGTGAGGCAGTGTCAGTAC
ATCTCGTCAACAAACAGCTAACAAGCGTCAGATGCATTCTTCGCAAAGCTGCAGGCAGCTTTGCAGCTTCACATCGACGATCCTGCCGTTCAAAGGCAGGTTCTGAGTTCTGGGGCGGATGAA
TTGGGCTTTGACTCTCTGGTTGCTGTTGAAATTCGGTCATGGTTCCTCAAAGAACTTGAAGTCGACATGCCTGTGCTTAAGATCCTTGGCGGGGCCTCTGTCTCAGACCTTCTTATTTTCGCCCT
TGATAAGCTGCCTGATGAACTAATCCCAAAAGTTGCGGCGATCCCTGAAGCTGTAGTAAACACGAGCAAAATCAAGAAAGATACGGTGACACCGCATTTCTCGCCATCAAGCGTTATTAGCTC
AAAAGCAACCAGTGATGGGGAAAGCCTCTTCTCGCCAGTCAATGGCTCTTCCCCTCCACAGAACTCTGTTGTCTCCCTTGTGTCTTCGTCGTCCTCTACGTATGACAAGGAAGAGCATTTCACC
CCGAATTATGAGCCTAAGTTTGAAAAGGTAGTGCCATTATCTTTTGGACAGTCACGATTTTGGTTCTTAAAGCATTATCATGAAGACCAGACAACGTTCAACATTACATTCTCGACTCGACTTAA
AGGTCCAGTGCGAATTCGTGACCTGGAAAACGCGGTGAGAACGATGGGCTATCGTCATGAGGCACTGCGAACGTGCTTTTTGTACAGGATTCCAACCAGCAGCCGATGCAGGGAATTCTAG
AGAAATCTTCCCTGTGTCTGGAGAAAATGAACATTCAGAATTCAAGTCAAGTCGCGGAAGAGTTTGAAGCAATGAAGAATCATGTTTTCGACATAGAGAATGGAGAAATTATGAGGGTTATT
CTCCTGTCCCTTAACCCGTTAAATAGCTTCCTCATTATCGGTTACCACCACATCAATATGGACGGTGCAAGCCTGGAAGTATTCCTGAGGGACCTTGAGTTGGTCTACTGTCGTAAACCACTGA
GTCAACCAGTATACCAGTATTCCGCTTTCTCGCTATGGCAAAGAAAGGAAATCGAATCCGGCAAAATGAAAGACGAGCTTCAGTACTGGAAGTCCGAATTTGCAGATCCTCCTACTCCATTGC
CATTGCTCCCTTTCTCCTCTACGAACAGACGCGAGGCAGTTTCAACATACAGTCATTATCGTGCCGATGTTCGAATTAATTCTCAGTTGGCAGAACAGATAAAGAATATGTGTTCGAAAAGGAA
AACCAGTATTTACCATTTTTATCTCTCCGTATATGAGGCAATGCTCTTCCGGCTCTTGGACACGAACGATCTCTGCATCGGCATGGCAGATGCCAACCGCAGCGAAGATACCTTTGTCAATAGT
ATGGGGATGTACTTGAACCTGCTTCCACTACGATTTCATCTTCGCTCGAATATGCAATTCATAGATGTGCTGAAAGAAACGCGCCGTAAAGCCTACAGTGCGATGGCTCATTCTCGTCTACCTT
TTGACGTTCTTCTTGACGAGCTGAAAGTACCACGCTCCACGTTGTATAGCCCATTGTTTCAAGCTTTTATCAACTACCGCCTGGGAGTCCAAGAAAAGAGAGCCTTCGGAGATCTCGAGTGTGA
GGGTGAGGAATACTCGTTAGGCAGAACTGCCTACGATATCAGTCTGGATATTCTTGATAACCAAAACAGCAATACCCTGCTTATGTTCATCGTACAAACGCAGTTATATTCGGCAGAGGATGC
AAGGCTTATGTGTAACATCTACGTAAGCCTTCTTGAACAGTTTTCCAAGGATCCAACCCTTCGTCTAGATGAGCCATCTCTCTTTGCAAAGGAGGATATTGCAGAGAATATTCAACTTGGAATG
GGTATAGTTTCCTATCTTCTAACGATAGAATGTGGCATTCACTAAAACAATCCTAATACCTTCTAGGACCGAATCACAACTCAGAATGGCCAGAGACCATGGTACACCGCGTCTACCAAATCAT
AGAGACATTCCCTGACAGTATTGCTCTTAAAGATACGGAGCACAACGAGTTGACATATCGACAGCTGGCTAATCGTATAGGCTGCATAGTATCATCTCTTCTTGAAGCAAATCTAGAAGAACA
TGCTCCTGTCGGTGTATTTCAAGAACCTTCACCTGATTGGGTCTGTTCAATTCTTGCTGTGATGGCTATTGGAGCCATCTACGTGCCATTTGACGCGAAATCGCCTACATCTCGTCTAGTGTCGA
TGATGGAAGATTGCCAGCCCACTGCTGTTCTCATCCATGATGCCACGGCCTCGAAGTTTGGAGGCTTGAGGTTGTCCTCTCGGGCAACTGTGATTAATCTTGCTAACATTCCATTAGACATCTC
GCACGAGAGTCGTAGTGTCGCAGCCAAAGCATGTGATCCTGCGGTGATACTGTACACCAGCGGTACGACAGGGGTACCGAAAGGGGTCATCTTGTCCCATGAGAGCCTAAGGAATGAGAAT
GAATTCGAGGTAGTCTCAGGTCCCGAAATTGTCCTTCAACAGAGTGCAATCAGTTTTGATCTCTCACTGAATCAAGTTTTCACGGCTCTGGCACATGGTGGAACACTTGTTATCATTCCAAGAT
CTATGCGGGGCGATGCAGTCGCAATAGCAAATATAATAGCCAAGGAAAAAGTCACCTATACTGGAGCAACGCCGTCAGAATATCTCAGTTGGCTTCAGTATGGGCGATCTGAGCTTTTGCAA
AACAAGGCTTGGAAGTTCGCCATGTCATGCGGAGAGCAGTACCCACAACAACTCGCTAATGAATTCAAACGCCTTGAACTTCCCCATCTAAGCCTGTGGAATGCATATGGACCAACCGAAGCC
ACTCTTTCGTCGAACAGAATCCAGTTATTCCTGAAAGATGATGATTCAAACCCACAGCATATTCCCGTTGGTTATACACTGACCAATTGTTCGGTATATATCGTAGACCGAGACCTCAATCCGTT
GCCTGCTGGAGTACCTGGGGAGGTCTGTATTGGGGGCGCCGGGGTCGCAATAGGCTACTTGAACAACAAGGAGCTTACCACAGAAAAGTTCGTGCCAAACCGGTTTGCAAATCCACATTTCC
TGTCGCGTGGGTGGCAACGAATGTATCGAACTGGCGACAGAGGAGTTCTCCAAGATGATGGAGCTCTGAAGATTCTGGGTCGCATTGACGGTGATACACAAATCAAGCTGCGTGGTATTCG
GATTGAAATGCAAGACATTGAGAATACCATTCTTCGTGAGGCAAACGGAGCGATATCGGATGTTGTCGTGGTGCCTCGTGGAGATCCTCCAGTCCTTATGGCTCATGCTGTCCTGTCATCCGT
CTCTCCGAAGGACCAGACCGGCTTTCTCCAGCATCTAAATACATCTCTTTCTTTACCACAATATATGCGGCCAGCAGCCATTATCCCTATCGATAAAATGCCTCTGACTACTCATGGCAAGATTG
ACAGAAGTGCGGTGCAGATCCTTTCGATAGCCCATCCCTCCCAAAAGCCATCTCCTGTCACTGAAATCACGCCAATGGAAAATCAACTTTGGCAGATCTGGGAGAATGTCATTTCAAAGGAAG
TTTTACAGTTGCACGCCATTGATGGTGATACAGACTTTTTCCATGTCGGAGGCAACTCGATGCTTCTCATCAAACTGCAAGAATTGATCAAACAGAAGTTTGGCACTTCGTTGGGTATCATGCG
TCTTTTTGAATGCAGCACGCTGCGTACGATGGCAGCAGCGATTCAAGACGCGTCCGTCACCAGGTCTCCTGATATTGTTCACTGGGAAGAAGAAATAGCTTTTCCCGAAGAGCTGTCAGTCAT
GGCCTCGTCCCAGTCTACTAACCTTACAAGAAAGCATGCTACTAGTCCTCTAGTAATCATCCTCACTGGTTCTACAGGCTTCCTGGGGAAAGAAATACTCCAACAGTTAATTGCATTGCCAAGT
GTCCAGGAAATTCACTGTGTTGCAGTCAGGGATGAGAAAAAGCTTGAGGACATGGTGACCTCCTCAGGCAAGGTAGTAATCCATCAAGGAGACTTGAGCCTTCCACAATGCGGGCTCTCCGA
TCAATCTGCCTCTGAAATTTTCAGCAAAGCAGACGCAATCATCCACAATGGCGCCGACGTGTCTTTCTTGAAGAGTTATCGATCCTTGAGCCCATCCAATGTCGAATCGACGAAAGAGCTGGTT
CGTCTGTCACTACGCCAACGGCAACTTATCCCATTTCACTTTATCTCAACCGTCACTACAGGACGACTGAATAAGAGCGAAACTTTTGGAGAAGTCTCGTTATCATCGACTCCTCCTCCTCCAGG
GTTCCCAGATGGTTACGTTGCCAGTAAATGGGCCAGCGAGGCCTTTTTGGAACGGGCAAGTACGAGTTTTAGTCTTCCGGTATACATTCACCGCCCATCGAGCATTACAGGCGAAGGAGCAA
GCGATCTCGATGTAATGCACAACATGCTAAAGTATGCACGCCTTACAAAATCCGTGCCTGAATCATCCCGCTGGAAAGGATATATCGACTTTATCAGTGTAGAGAATGTCGCTCGCGGGATCA
TCCAGGAAGTCCTCCACAGGGAAAACGAAAGTAGTACTCCCGCAAGCAATACCGCTCATGTGAGATACATTCATCATTCCGGAGATATCGTTACCACAGTCGAGTCGTTTAAAGACTACTTGA
CAAATGAAAGTGGAGAAACCTTTATAGGATTGCCGCTAAGAGAGTGGGTGGAAAGGGCAAAGGAACACGGCCTGAATGTCCTAGTCGCGGGTTACTTGACAGCAAGTGATGAGAGACAGG
TGGATGTGGTATTCCAGAGGTTGATTAAATTGAAAGTGGCATAAGGAATTCGTAGATCGGGGCAGAGCGGTTTCTTGAATTTGCTTATTCTCAATGGAGATTGAACATTGATATGACGAGCTT
TACTATTTCCATCATCTTGAAAAAATGCAAGGTCAGAAAGAATAAGTTTATTTTTCTGTGCCAAAGGATATCCGGGCCTGCTACTTACGGCCAATTCTACAGTAGAGACCGGCCAATTCTGGCG
CCCATGAAATGTAAATATGTGATATTGGTTGCAGATCTTGCCATATCTGTTATCAAAATATCAGGGCGGACTTTATATTTTTATTTTAAATATAGCTATCACCCAAATTCCATGTAATTATGTAAT
ATACTGGACATCTCTGGCTTAAAAGAGCTAGTATTAATACCTAAAACTCTAAGCTCTGGTTCTTGAGTCCTGCCTCTCTCTAGGTTTTGTAATTCCAAAATCACCTTTCGTGCCCAGTCAAGGTCT
ATTATTTTGTACTGGGTTGGATCACTGGGGATAACCTCACGGAGTCTGTTCCAGACAGAAAGAGGCCAGGCACGCTTTTGCATATTCGAGAGACCATGGAACCAAGATTTCCAAGAGCTTTCT

CTCCGGGAAAATAGAAATTAAGAGCGGCCATAGCGATTCTGGTAGGATGATATTATCGTTTTTCTCCAACAGAAGTCTCAATTTGACGGCGGGCTCGGGTAAAGATAGTGGCCTCGTTATTGG
CCGGGAGATAGCCACAGAGGATCTGGATCCGGAAGACTCGCTGATGAATGAGCTTGCGAGAAGTGTCATGGGCGAGATAGAGGGGTTTGGGTGCTAATTCAAATCTCGTAGATTGGGTACT
TGTTCCAGGATATCAAGTGGGATACGGGAGTGCACTTTGTACCAATCACACCGACTGTGGAAGGAGATGATATAACAACTTTAGGTGACGGGCCTAGGCTGCAAAGCATTTGCTCCGTACCT
AATTAAAAAAAAATGTAGCAAGTTTGATATTATACTATGTATTGTAAGATTGTGCGCGCTAAGAAAAGTGAGGAGGTTTGTATACTATCGACTAAAGCTCGCTATCCAACAAGAGCACATGAGT
GGATTTATGCGCTGCAATCGCCATATTGGGATAGCCGTCAGGTAATCTGGTCTCTAAGCTAAGTTATTCCAAAACTAATATCATTTTGCCAAGTCAATCCTTCTTGAATCGACTTGGGTCCATGC
ATATAATGATTGGTTCCCATGTCTGACCCCACAGGCTGATGCAAATGTAACGACGCGTTAACCAAGCGACAAAACATCGGAACAACGTACTAGGTAGCTGGGAATGTACCTGCCACAGAATTT
TGCGCAATCTCAGGTAAGGTCGGAATGCACTGGGATGCCCTCAGTTAGTTAATGCCCCGGGTATATTAGGGCCGAAATGCAGACAACGGTGGACGATGAACAAACTGTAGGCGCTCACAAG
TTTCTTGTCTTTGTGGAACGAAATCGACACTCACTTCACTTGCTTGATCGCGACCGGGCCTATTTGATCGCTCAATTGGATCAAATGTCAGCCAACCAATTATCTTGGCCAAAGAGTGCCGCAG
ACGTGTCAGACATGTTCGTTTTCATGCTGCAAGCACAGTGGATATCCGCCCAATACCATGGCCGGCACCCAGACTCAGATTTTGCAGATAAAGAATTGTGGAGAATACGGGAGAATCGAGGA
GTCAATTCTCAGGGAAAGCTTCGAGAGTATGCAAAGGAGCGTAACATGGATTATAATAGGCTTCGGGATACGGTTCAAGTTAACCAGAAAGCGGGAGCCATAGAGGCGGAGTTGGGAGTTC
CCGGCTTGGGAAAGACCTTATACTCTGCCTTCAAACGACTAAGAGATCTGAAGCAAAATGAAATTGAGTCATTCTGCAGGCTATTTCGTGCTAGTGATTCGTTTGCAGATTACAAGAGCTATG
CAGCTATAGCGCTGGCATTATTGGACGATTATTTGACTCTGATGCAGCGGCATGATTTAGATCAATAGATAGTGGGATAAAATGACAAACTATGCATGCTTATCGAGAAGTTCTCACCCCTTTT
CCAGGTTTCTGATCGTAGGTGTTGCATGGAATGGCGCTACTTTGACGGGGACATCATCCCTCTGAATCATAATCCGTTGCCGTTCAAGAGAAAAGGTCTGGAGAACAAATAGACCATCACTTT
CACTCGGCCAATTGAAGCACATATCTGATAGTATATAGGCTACCACTTTCTGAGATAACTCTAGCGGTACCAGGCATAGAGGACCTTCTAATCTCTTCCGATCGAGAGTTGTCCGGTCGAGCGT
ATGCGACTTATCGTATAGCTGAAACCGAACGTTTCCCCGCACAAGCCGCAGCAAGCAGATTGCTTTGTTTCGATGGTATCCGTAGGGTCCTAGCCTATAGATTGAGGGCGGTTCTGTGGCTTG
GGCGCAGTTCACTAGAATGGGTTCTCCACGAACGACATCCTTTATGTCATTTCCACCAAGAAGGGAAAGAATGTAGTGCATCTATCAGCTCGGCGCGTAAGCTTGTGGAGTATATGTTTTTCA
CGAATAACGTACCCTTGCTTCCAATAGGTCATGCTCTACGAAGTTAACGAGAGCTTTCGATAATTGAGTCCCGTATCTGCCCCGGACACAAAGTGTGGTTGTCGAGTTCTGTTCGAGAACATCA
TCCAGCCACTTCAGTGCCTGGCCCGACGGAGGGTCCCAAGGCAAAACAGCATAACCTTCGTAGTTGAAAGTCATCTCCTCATCTTGTGCGTCATCCTCGGGAGCGCTCACAACGGCCCCGCGA
TATTTCCTCGTTGCATTCTACAGTAGTGTATGATGATTAGGTAATCATCCTTTGGCCGTGGGTTCTGAAATCAAGGCGTAGATAGGCCGCATGCGCAAGTGATGAGAGATGTGATAGTCACAG
AACGTGTTTATTCTCGAATCAAAGTTTGGGACTGAGTGGAAATTGACTTAATCTCTCTCATAGGTGCTCAAGGCTAGAATATATGAA

S2 Fig. Deletion of the green pigment UA08_00425 (*albA*) in *T. atroroseus*. Second independent trial. A-C) Plates resulting from co-transformation of pD-hyg-UA08_00425 and CRISPR-Cas9 vectors carrying three different protospacers, protospacer 1-3, respectively. D) *T. atroroseus* transformed with gene-targeting plasmid pD-hyg-UA08_00425 (pFC574) in the absence of a CRISPR-Cas9 vector.



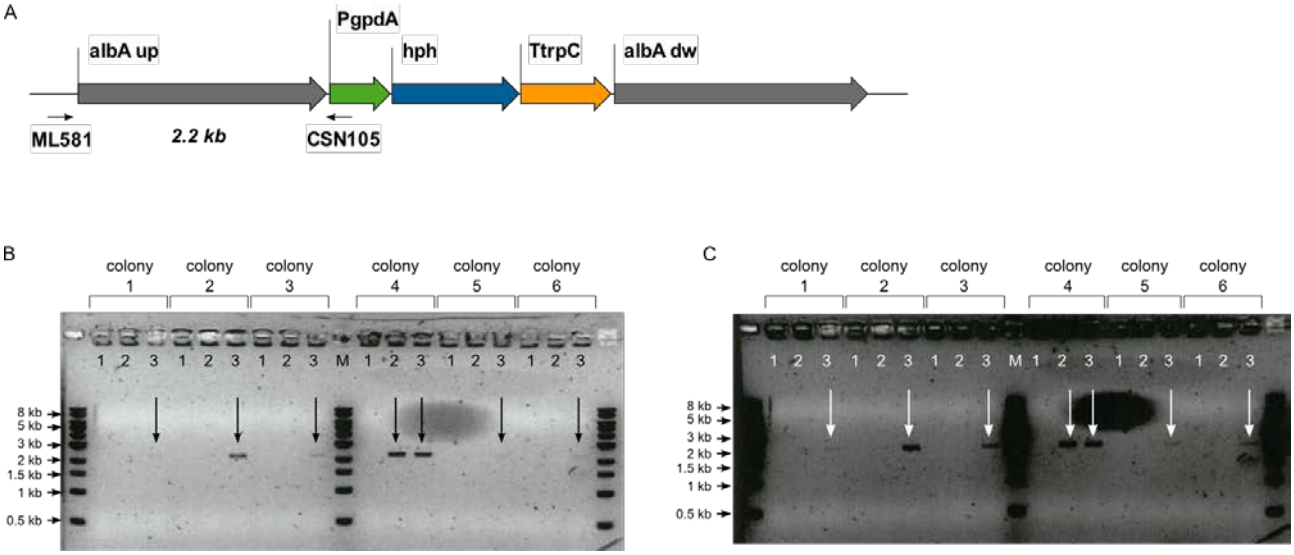
S2 Table. List of plasmids

pFC330	pCas9-pyrG	Empty CRISPR vector
pFC332	pCas9-hph	Empty CRISPR vector
pFC476	pD-hyg	Deletion vector with *hph* marker
pFC574	pAMA1-hph	AMA1 *hph* vector
pFC683	pCas9-pyrG-talA	CRISPR plasmid for deletion of *talA*
pFC687	pD-hyg-talA	*talA* deletion plasmid
pFC784	pCas9-pyrG-UA08_00425-PS1	CRISPR plasmid for deletion of *albA*
pFC785	pCas9-pyrG-UA08_00425-PS2	CRISPR plasmid for deletion of *albA*
pFC786	pCas9-pyrG-UA08_00425-PS3	CRISPR plasmid for deletion of *albA*
pFC789	pD-hyg-UA08_00425	*albA* deletion plasmid

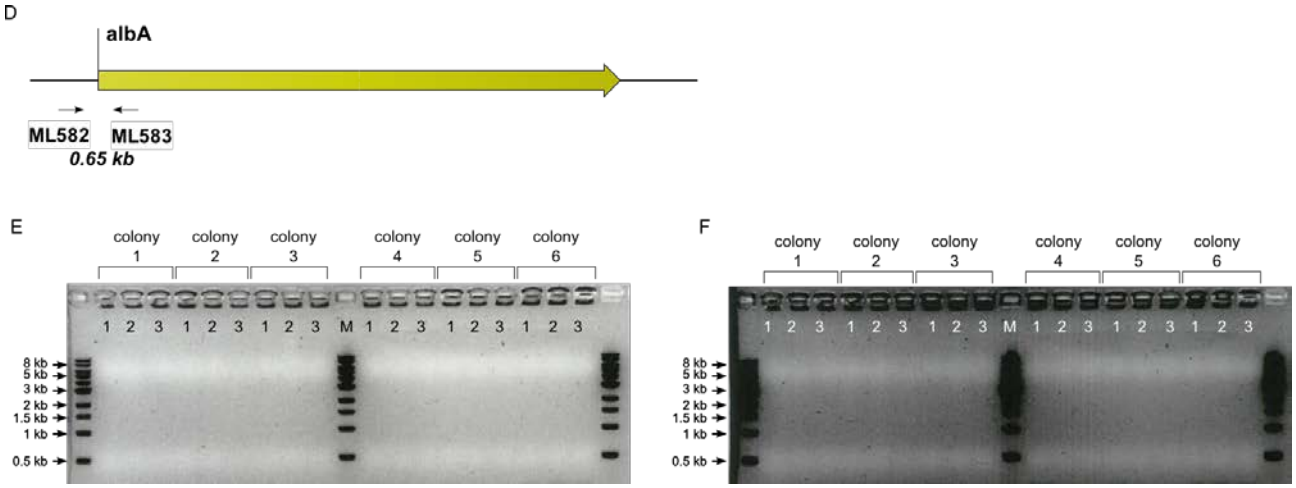
S3 Fig. Tissue PCR analysis for verification of *albA* deletion.

The protocol for tissue PCR is described by Nødvig *et al*. For each PCR reaction, a dilution series of *T. atroroseus* mycelium was done to achieve optimal DNA template concentrations in one or more reactions. For each transformant, three PCR reactions were performed by adding approximately 1 mm^2 mycelium template to three tubes by sequentially dipping the pipette tip with the biomass in each tube. Hence, the concentration of biomass was highest in the first tube and lower in the following two tubes. The resulting PCR reactions were loaded on an 1% agarose gel as sample 1, 2 and 3, respectively, as indicated on the gel picture. Lane labeled M includes the 1 kb ladder from New England Biolabs.

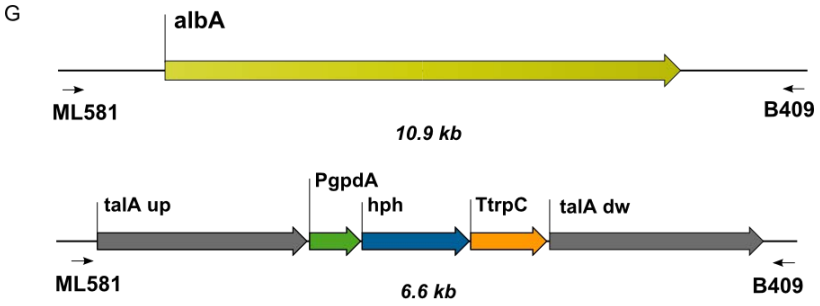
Three setups were performed for analysis of transformants to check for deletions of the green conidia pigment gene *albA* (UA08_00425). In the first PCR setup (A), the forward primer (ML581), which binds outside of the upstream *albA* targeting sequence, and the reverse primer (CSN105), which binds to the P*gpdA* promoter of the *hph* marker, are used for PCR. A successful gene replacement of *albA* with *hph* results in a PCR band of ~2.2 kb. Analyses of six randomly selected white transformants (corresponding gel lanes are indicated as colonies 1-6) using the first PCR setup are shown in (B) and (C). The number of transformants is indicated above each gel. (A) and (B) depict the same gel, but with different exposure times, hence allowing very weak bands to be visualized. Arrows point to successful PCR reactions.

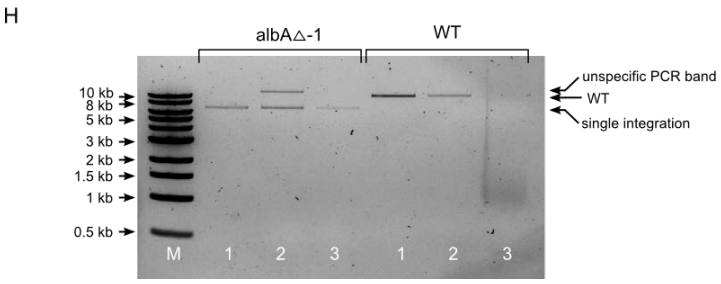


In the second PCR setup (D), the upstream primer (ML582) binds immediately upstream of the *albA* start codon, while the reverse primer (ML583) binds inside the *albA* ORF. The presence of a wild-type *albA* locus results in a band of ~0.65 kb. Analyses of six randomly selected white transformants using the second PCR setup are shown in (E) and (F). The number of transformants is indicated above each gel. (E) and (F) show the same gel, but with different exposure times. No visible bands can be detected indicating complete deletion of *albA*.



In the third PCR setup (G), the forward primer ML581 and the reverse primer B409 are used to amplify the *albA* locus. Both primers are situated outside the regions used as homology sequences of the *albA* gene-targeting substrate). The wild-type *albA* locus (upper cartoon) produces a band of 10.9 kb (upper cartoon) whereas a successful gene replacement of *albA* with *hph* (lower cartoon) results in a band of 6.6 kb. Analyses of a randomly selected white transformant and a wild-type strain (indicated as *albAΔ-1* and WT, respectively) using the third PCR setup are shown in (H). Arrows indicate the positions of the bands expected from wild-type *albA* (6.6 kb) and *albAΔ::hph* (10.7 kb). The position of an unknown unspecific PCR band of unknown origin is also indicated.

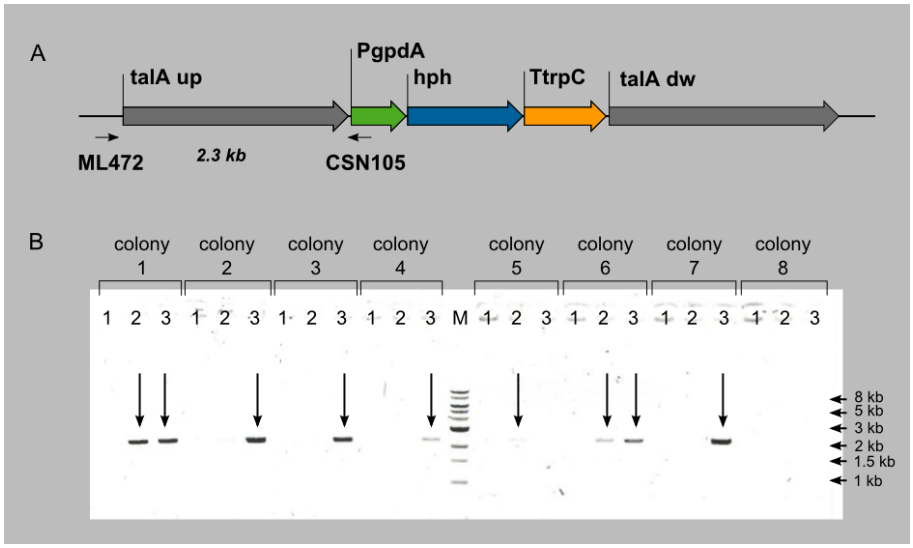




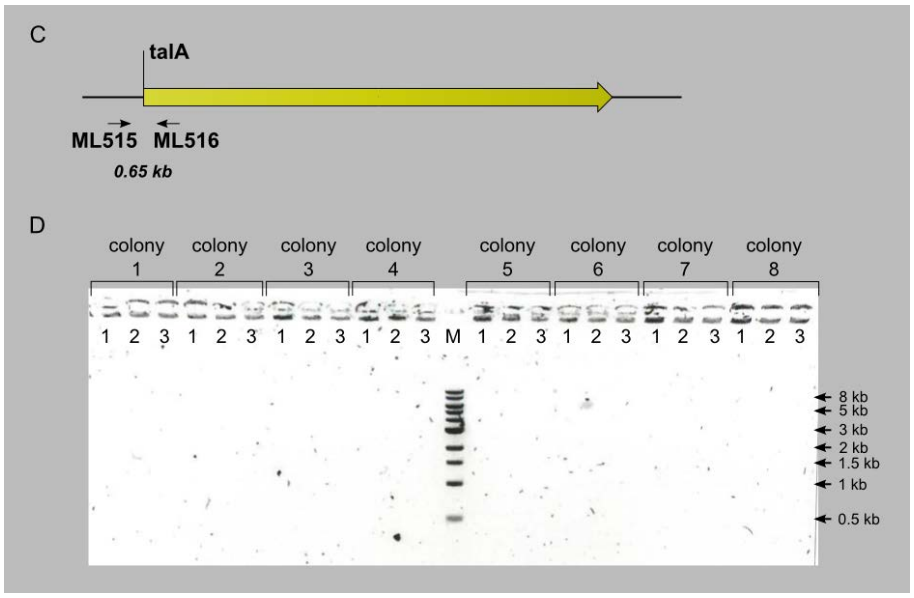
S4 Fig. Tissue PCR analysis for verification of *talA* deletion.

The protocol for tissue PCR is described by Nødvig *et al*. For each PCR reaction, a dilution series of *T. atroroseus* mycelium was done to achieve optimal DNA template concentrations in one or more reactions. For each transformant, three PCR reactions were performed by adding approximately 1 mm^2 mycelium template to three tubes by sequentially dipping the pipette tip with the biomass in each tube. Hence, the concentration of biomass was highest in the first tube and lower in the following two tubes. The resulting PCR reactions were loaded on an 1% agarose gel as sample 1, 2 and 3, respectively, as indicated on the gel picture. Lane labeled M includes the 1 kb ladder from New England Biolabs.

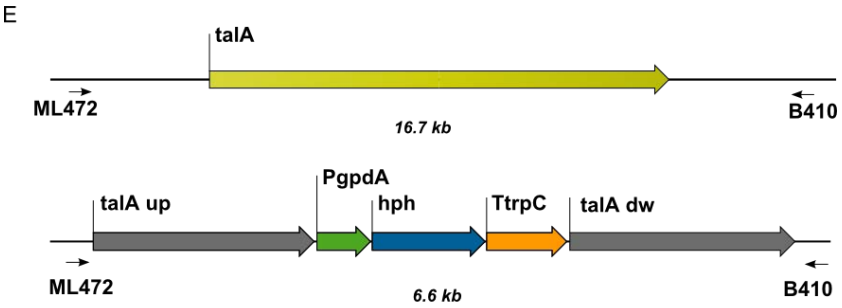
Three setups were performed for analysis of transformants to check for deletions of *talA* (UA08_04452). In the first PCR setup (A), the forward primer (ML472), which binds outside of the upstream *talA* targeting sequence, and the reverse primer (CSN105), which binds to the P*gpdA* promoter of the *hph* marker, are used for PCR. A successful gene replacement of *talA* with *hph* results in a PCR band of ~2.3 kb. Analyses of eight randomly selected transformants (corresponding gel lanes are indicated as colonies 1-8) using the first PCR setup are shown in (B). The number of transformants is indicated above each gel. Arrows point to successful PCR reactions.

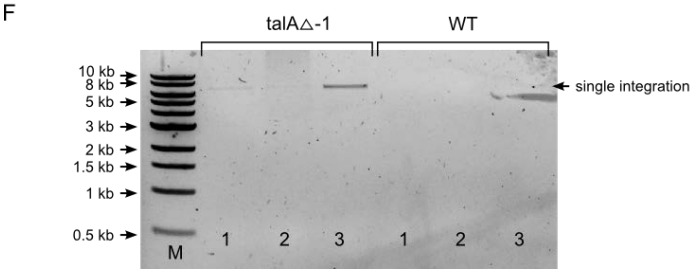


In the second PCR setup (C), the upstream primer (ML515) binds immediately upstream of the *talA* start codon, while the reverse primer (ML516) binds inside the *talA* ORF. The presence of a wild-type *talA* locus results in a band of ~0.65 kb. Analyses of eight randomly selected transformants using the second PCR setup are shown in (D). The number of transformants is indicated above each gel. No visible bands can be detected indicating complete deletion of *talA*.



In the third PCR setup (E), the forward primer ML472 and the reverse primer B410 are used to amplify the *talA* locus. Both primers are situated outside the regions used as homology sequences of the *talA* gene-targeting substrate). The wild-type *talA* locus (upper cartoon) produces a band of 16.7 kb (upper cartoon) whereas a successful gene replacement of *talA* with *hph* (lower cartoon) results in a band of 6.6 kb. Analyses of a randomly selected transformant and a wild-type strain (indicated as *talAΔ-1* and WT, respectively) using the third PCR setup are shown in (F). Arrows indicate the positions of the bands expected from wild-type *talA* (6.6 kb) and *talAΔ::hph* (16.7 kb). Note, that while the *talAΔ*-1 allele can be detected by PCR; amplification of the entire *talA* locus in the wild-type strain was unsuccessful likely due to the size of the PCR fragment.

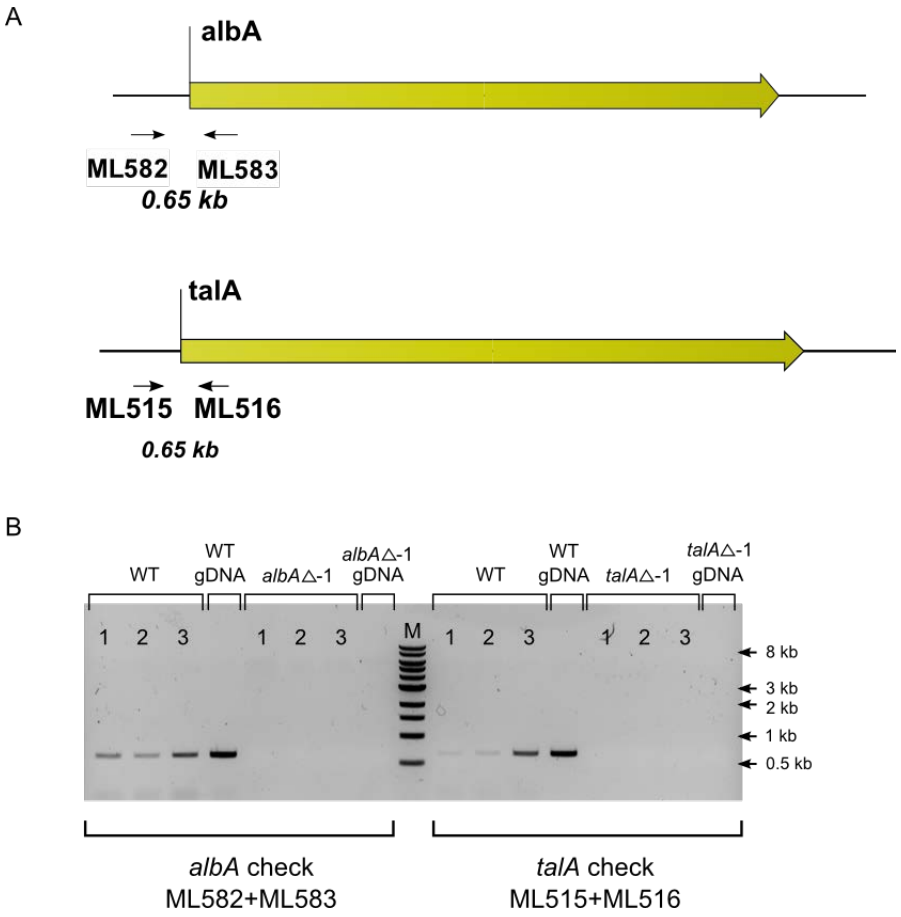




S5 Fig. Tissue PCR analysis for verification of primer functionality.

The protocol for tissue PCR is described by Nødvig *et al*. For each PCR reaction, a dilution series of *T. atroroseus* mycelium was done to achieve optimal DNA template concentrations in one or more reactions. For each transformant, three PCR reactions were performed by adding approximately 1 mm^2 mycelium template to three tubes by sequentially dipping the pipette tip with the biomass in each tube. Hence, the concentration of biomass was highest in the first tube and lower in the following two tubes. The resulting PCR reactions were loaded on an 1% agarose gel as sample 1, 2 and 3, respectively, as indicated on the gel picture. A PCR reaction using purified genomic DNA as template was also included in the setup. Lane labeled M includes the 1 kb ladder from New England Biolabs.

The functionality of the primers used in S3 Fig (D-F) and S4 Fig (C-D) were tested using wild-type (WT) *T. atroroseus* tissue and gDNA. Using the primer sets for the *albA* locus (ML582+ML583) and the *talA* locus (ML515+ML516) results in a PCR band of ~0.65 kb in the WT strain in both loci. (A) shows positioning of primers at the *alba* and *talA* loci. (B) Agarose gel electrophoresis (1 %) of PCR samples obtained from WT *T. atroroseus* as well as from a single transformant of each of *albA⊿* and *talA⊿* are shown in (B).



RESEARCH ARTICLE

Linker Flexibility Facilitates Module Exchange in Fungal Hybrid PKS-NRPS Engineering

Maria Lund Nielsen, Thomas Isbrandt, Lene Maj Petersen¤, Uffe Hasbro Mortensen, Mikael Rørdam Andersen, Jakob Blæsbjerg Hoof, Thomas Ostenfeld Larsen*

Department of Systems Biology, Technical University of Denmark, Søltofts Plads, Kongens Lyngby, Denmark

¤ Current address: Lundbeck A/S, Ottiliavej 9, Valby, Denmark
* tol@bio.dtu.dk

OPEN ACCESS

Citation: Nielsen ML, Isbrandt T, Petersen LM, Mortensen UH, Andersen MR, Hoof JB, et al. (2016) Linker Flexibility Facilitates Module Exchange in Fungal Hybrid PKS-NRPS Engineering. PLoS ONE 11(8): e0161199. doi:10.1371/journal.pone.0161199

Editor: Richard A Wilson, University of Nebraska-Lincoln, UNITED STATES

Received: June 29, 2016

Accepted: August 1, 2016

Published: August 23, 2016

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by the Novo Nordisk Foundation (http://www.novonordiskfonden.dk/en). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Polyketide synthases (PKSs) and nonribosomal peptide synthetases (NRPSs) each give rise to a vast array of complex bioactive molecules with further complexity added by the existence of natural PKS-NRPS fusions. Rational genetic engineering for the production of natural product derivatives is desirable for the purpose of incorporating new functionalities into pre-existing molecules, or for optimization of known bioactivities. We sought to expand the range of natural product diversity by combining modules of PKS-NRPS hybrids from different hosts, hereby producing novel synthetic natural products. We succeeded in the construction of a functional cross-species chimeric PKS-NRPS expressed in *Aspergillus nidulans*. Module swapping of the two PKS-NRPS natural hybrids CcsA from *Aspergillus clavatus* involved in the biosynthesis of cytochalasin E and related Syn2 from rice plant pathogen *Magnaporthe oryzae* lead to production of novel hybrid products, demonstrating that the rational re-design of these fungal natural product enzymes is feasible. We also report the structure of four novel pseudo pre-cytochalasin intermediates, niduclavin and niduporthin along with the chimeric compounds niduchimaeralin A and B, all indicating that PKS-NRPS activity alone is insufficient for proper assembly of the cytochalasin core structure. Future success in the field of biocombinatorial synthesis of hybrid polyketide-nonribosomal peptides relies on the understanding of the fundamental mechanisms of intermodular polyketide chain transfer. Therefore, we expressed several PKS-NRPS linker-modified variants. Intriguingly, the linker anatomy is less complex than expected, as these variants displayed great tolerance with regards to content and length, showing a hitherto unreported flexibility in PKS-NRPS hybrids, with great potential for synthetic biology-driven biocombinatorial chemistry.

Introduction

Polyketide synthases (PKSs) and nonribosomal synthetases (NRPSs) are among the major biosynthetic enzymes for fungal secondary metabolites, and are responsible for the biosynthesis of numerous medically relevant compounds including statins, mycophenolic acid, cyclosporine, and

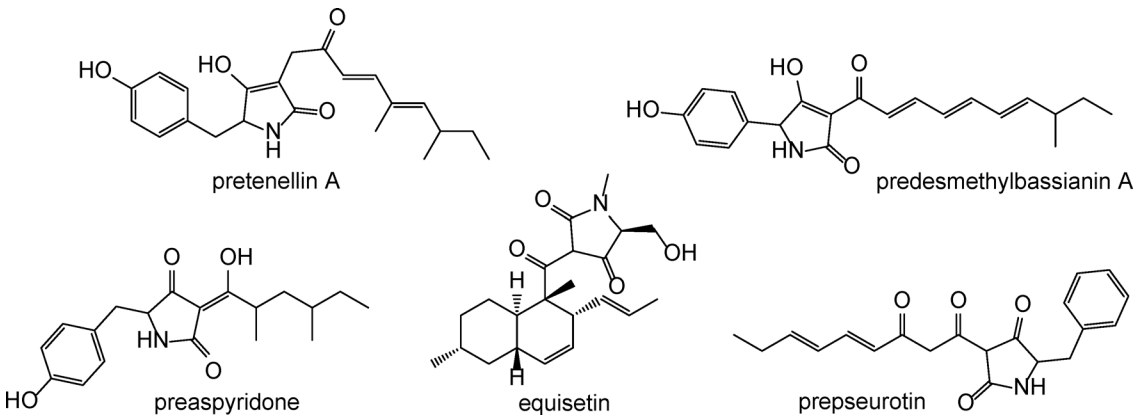


Fig 1. Structures of known PKS-NRPS products [5–9].

doi:10.1371/journal.pone.0161199.g001

penicillin [1,2], and within the last ten years, natural fusions of PKSs and NRPSs (PKS-NRPSs) have been described [3,4]. These large modular enzymes consist of a type I iterative highly reducing PKS fused to a single NRPS module. Characteristic of fungal PKS-NRPSs is the lack of a functional enoyl reductase (ER) domain, and most of these enzymes therefore rely on a *trans*-acting ER for production of polyketide-amino acid compounds [4]. Most commonly, the PKS-NRPS hybrid and its cognate ERs are encoded in the same cluster. From several studies, the products of PKS-NRPS hybrids co-expressed with their cognate ER have been revealed (Fig 1).

Fungal PKS-NRPSs have been reported to produce highly bioactive compounds, with some of the most prominent examples being pseurotin, cyclopiazonic acid, fusarin C, and cytochalasins [10–13]. For many years, the most common approach to discover new natural drug leads has relied on screening of a large number of organisms often followed by semi-synthetic modifications for final drug structure optimization [14]. For a more sustainable and economical production of natural product derivatives, synthetic biology offers new alternatives, and recent discoveries have sparked the interest for producing analogs of polyketides, nonribosomal peptides, as well as hybrid compounds using combinatorial biosynthetic approaches. By exploiting the modularity of enzymes involved in secondary metabolism, it has been proved feasible to produce novel synthetic compounds by identifying genetic modules, and combine them as building blocks at the genetic level, the goal being the rational design of novel chimeric proteins with desired catalytic properties [15].

The limited structural information on fungal iterative PKSs is the biggest obstacle towards understanding enzyme programming, and thus, product formation. Several studies have been conducted on natural hybrid-, chimeric- and dissected PKS-NRPSs and successful construction of functional chimeric PKS-NRPSs has been achieved in a few studies. One example is the elegant domain and module swapping involving several PKS-NRPS variants, where Cox and coworkers were able to resurrect the extinct metabolite bassianin, as well as reveal some of the underlying mechanisms for product formation [5,16]. The two enzymes used for these experiments, TenS and DmbS, are involved in the production of tenellin and desmethylbassianin, respectively, and the genes encoding TenS and DmbS (87% sequence identity) both originate from the insect pathogen *Beauveria bassiana*. In another study [8], the PKS module of the aspyridone producing PKS-NRPS (ApdA) and the NRPS module of the PKS-NRPS involved in production of cyclopiazonic acid (CpaS) were expressed as individual proteins in *Saccharomyces cerevisiae*, which led to the incorporation of a tryptophan residue into the aspyridone polyketide backbone. Most of the studies have considered only single module swaps and it has

therefore been difficult to determine general mechanistic similarities. Recently, a comprehensive study of PKS-to-NRPS compatibility was conducted by Schmidt and co-workers [9]. They constructed 34 distinct module swaps, and in addition to revealing compelling new information on the programming rules of hybrid PKS-NRPSs, they succeeded in the production of a chimeric PKS-NRPS product. Fusion of the equisetin PKS module (EqiS) with the fusaridione A NRPS module (FsdS), both from *Fusarium heterosporum*, resulted in production of the predicted chimeric compound.

So far, no studies have investigated the importance of the PKS-NRPS inter-modular linker. From the existence of many diverse modular proteins in nature, it has long been known that gene duplications as well as the modular assembly of existing genes is a major source of evolutionary novelty [17]. Multidomain proteins are thought to have evolved by gene duplications or by shuffling of sequences encoding different protein domains. From studying the sequence, it is evident that protein domains and modules are often separated by linker sequences that vary greatly in size, and it is well-known that the properties of these linkers are highly sequence-dependent. Changes in the length and flexibility of the linker can have several implications for protein stability- and folding rates, domain-domain interactions, and enzyme activity [18]. Despite the interest in engineering of compounds of mixed biosynthetic origin in fungi, no studies have thoroughly looked into the role of the inter-modular linker of fungal PKS-NRPS hybrids.

In this work, we chose to apply synthetic biology to cytochalasans due to their wide range of distinctive biological functions [19]. We report successful module swapping between the two PKS-NRPS hybrids CcsA from *A. clavatus*, which has been shown to be involved in the biosynthesis of cytochalasin E [20], and the related previously uncharacterized Syn2 from the rice plant pathogen *Magnaporthe oryzae* leading to novel hybrid products heterologously expressed in *A. nidulans*. For the first time, we have methodically tested the inter-modular PKS-NRPS linker for its role in the transfer of biosynthetic intermediates. By expression of several linker-modified variants, we demonstrate that these linkers display great tolerance with regards to content and length, showing a hitherto unreported flexibility in PKS-NRPS hybrids, with great potential for synthetic biology-driven biocombinatorial chemistry.

Results and Discussion

Co-expression of *ccsA* and *ccsC* in *A. nidulans* leads to production of a modified cytochalasin intermediate

To investigate if the *A. clavatus* hybrid PKS-NRPS (ACLA_078660) could be functionally expressed in *A. nidulans*, *ccsA* along with the *trans*-acting ER *ccsC* (ACLA_078700) encoded in the same gene cluster, were transformed in a two-step approach into *A. nidulans*. The strain was analyzed by ultra-high performance liquid chromatography (UHPLC) coupled with diode array detection (DAD) and high-resolution mass spectrometry (HRMS), and the metabolite profile revealed the appearance of a new major compound with a mass of 415.2585 Da ($[M+H]^+$ = 416.2584) corresponding to the elemental composition $C_{28}H_{33}NO_2$. It was found that detection of this product was completely dependent on co-expression with the ER as no products were detected in its absence. We successfully purified the heterologous product of the CcsA/CcsC expressing strain, and NMR structural elucidation (see S1 Dataset) revealed a hybrid polyketide-nonribosomal peptide, consisting of the expected phenylalanine moiety joined to a decalin scaffold, originating from a highly reduced polyketide chain, via a tetramic acid derived lactam (Fig 2). We named this new heterologously expressed hybrid product niduclavin.

Considering the general structure of cytochalasins such as cytochalasin E (Fig 3), two elements of the structure of niduclavin were unexpected. A double bond was found between the C-2' and C-3' position of the phenylalanine side chain, which is a feature that to our knowledge

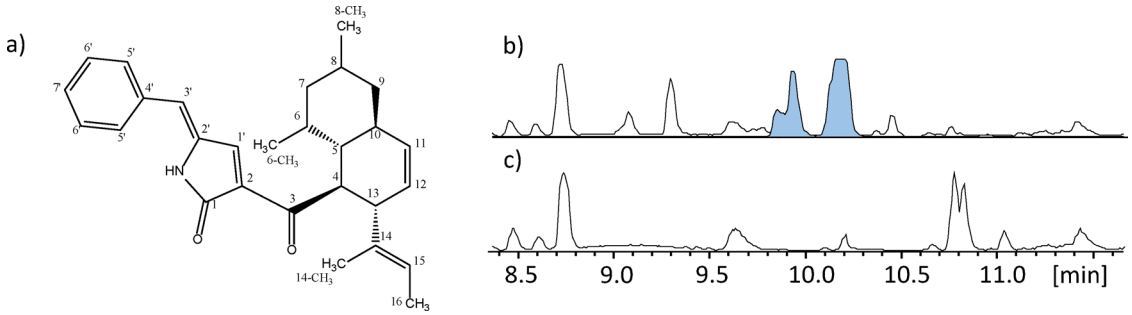


Fig 2. Overexpression of *ccsA* and *ccsC* in *A. nidulans* leads to production of niduclavin. A) The structure of niduclavin, elucidated by NMR spectrocopy, B) Base peak chromatogram (BPC) of *A. nidulans* extracts, showing production of niduclavin (extracted ion chromatogram (EIC) @ *m/z* 416.2584 highlighted in blue), and c) BPC of reference strain, which displayed no production of niduclavin.

doi:10.1371/journal.pone.0161199.g002

has only been found in talaroconvolutin A, from *Talaromyces convolutes* [21] and myceliothermophin E, from *Myceliophthora thermophile* [22] (Fig 3). Even more interesting, the [4+2]-cyclisation normally encountered in cytochalasin biosynthesis is absent. Instead, a decalin ring system was found, rather than the normally observed 11 membered macrocycle fused to a bicyclic lactam (isoindolone).

We speculate that cross-chemical reactions with endogenous *A. nidulans* activities are responsible for the introduction of the double bond. By introducing the additional double bond, additional activation of the dienophile in the α/β-position of the ketone (C-3) would occur, thereby possibly favoring the decalin formation in this position rather than at the tetramic acid moiety present in the *A. clavatus* molecule (see S1 Fig). Whether this reaction requires an enzymatic activity is unclear, however the formation of a couple of earlier eluting likely isomeric niduclavin analogues (Fig 2B), suggests that the reaction is non-enzymatic. In nature, there are numerous examples of PKS- and PKS-NRPS products containing decalin ring systems, *e.g.* lovastatin, equisetin, talaroconvolutin, and codinaeopsin [21,23–25]. Since the decalin ring system is not present in the cytochalasins, it seems likely that formation of the isoindolone moiety in *A. clavatus* is enzyme catalyzed.

We searched the literature for possible candidate enzymes in *A. nidulans* that could be responsible for the introduction of the double bond between the C-2' and C-3' position of the phenylalanine side chain. The dioxygenase AsqJ (ANID_09227) from *A. nidulans* has previously been shown to introduce a double bond in the same position of a phenylalanine moiety of an intermediate in the cyclopenin/4'-methoxycyclopenin biosynthetic pathway [26]. However, deletion of *asqJ* did not result in any changes of the final product of CcsA and CcsC. As an alternative approach to create a proper cytochalasin intermediate, the two genes were also

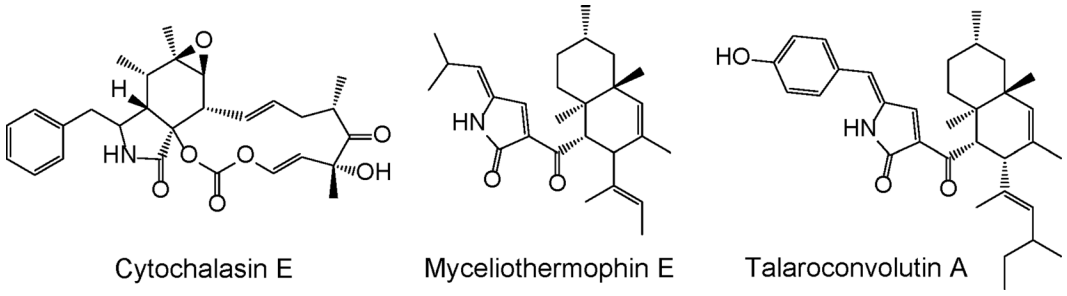


Fig 3. Structures of cytochalasin E, myceliothermophin E, and talaroconvolutin A.

doi:10.1371/journal.pone.0161199.g003

expressed in *A*. *niger*. Yet, expression of *ccsA* and *ccsC* in *A*. *niger* also lead to production of niduclavin (data not shown). The results point to the potential problems with heterologous expression, meaning we cannot by default assume that a product or intermediate produced in a non-native host is identical to the compound produced in the native organism.

In addition to niduclavin, the extract of the CcsA/CcsC expressing strain contained four additional compounds, including one with the mass 440.3188 Da. This mass is identical to the mass of the cytochalasin precursor product that was described by Fujii *el al*. [27] in *A*. *oryzae*, and we speculate that our compound is identical to theirs. Our results and the results of Fujii *et al*. suggest that formation of the characteristic cytochalasin isoindolone and macrocycle moieties in *A*. *clavatus* requires a chemoselective functionality in order to direct the rearrangement reaction towards the macrocyclic pre-cytochalasin product. The existence of this type of activity is supported by a study by Kasahara *et al*. where the biosynthesis of solanapyrone from *Alternaria solani* was investigated [28]. They identified a flavin-dependent oxidase catalyzing an oxidation and mediating a cycloaddition. To our knowledge, no equivalent activities have been found in the *A*. *clavatus* genome. Recently, Klas *et al*. [29] called into question the existence of true Diels Alderases, and it appears that enzymes catalyzing [4+2]-cycloadditions in most cases serve as multifunctional enzymes, *e*.*g*. oxidations as seen in the case of solanapyrone biosynthesis. If the [4+2]-cycloaddition in cytochalasin E is a secondary activity of another enzyme, it is conceivable that the formation of the cytochalasin isoindolone moiety is mediated by one of the tailoring enzymes encoded in the *ccs* cluster. However, cytochalasin biosynthesis in *A*. *nidulans* is believed to be hampered by the formation of the decalin ring system, hereby preventing the formation of the expected isoindolone moiety.

The Syn2 PKS-NRPS hybrid produces a novel polyketide-nonribosomal peptide intermediate

Based on previous studies, we hypothesized that construction of active chimeric PKS-NRPSs would be dependent on the degree of sequence identity between the recombined modules [5,9,16]. A BLAST search identified a putative PKS-NRPS from *Magnaporthe oryzae* (CAG_28798) with 68% amino acid sequence similarity as the closest homolog to *A*. *clavatus* CcsA (52% identity). The gene, known as *syn2*, is encoded in a previously described gene cluster where another PKS-NRPS hybrid (*ace1*) is also found [30]. Associated genes encoding ERs are predicted for both *ace1* and *syn2* (*rap1* and *rap2*, respectively). Special attention has previously been devoted to this gene cluster because it was shown that the product of the Ace1 pathway is an avirulence factor and is recognized in *Pi33* rice cultivars making them resistant to fungal infection [31]. It was later demonstrated that genes of the Ace1 gene cluster were expressed exclusively in the appressorium during infection of the host plant, although, deletion of *syn2* did not affect avirulence [30].

The only *M*. *oryzae* strain having a publically available genome sequence is the laboratory strain 70–15 [32], which is derived from the wild-type strain Guy-11 through several backcrossings [33]. However, it has been reported that the *syn2* allele in strain 70–15 is inactive, due to an early stop codon from an insertion of a single base pair [30]. The sequence of *syn2* from the wild-type strain Guy-11 has been published (CAG28798) [31], and this allele does not contain this stop codon [30]. To analyze the function of Syn2, we purchased the coding sequences of the Guy-11 *syn2* and its corresponding ER *rap2* (MGG_08380). Seven introns are annotated in the publicly available sequence of Guy-11 *syn2* although only five are predicted using the Augustus gene prediction software (http://bioinf.uni-greifswald.de/augustus/). By including the two putative introns (nucleotides 7962 to 8082 and 8145 to 8323) as part of the coding sequence, a complete amino acid alignment of *syn2* and *ccsA* was possible. To support both

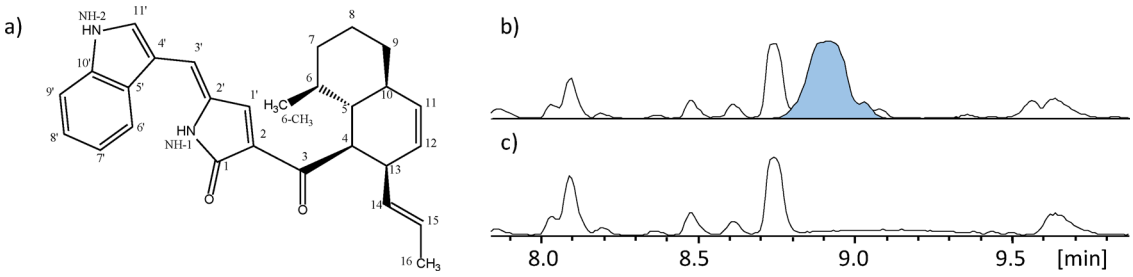


Fig 4. Overexpression of *syn2* and *rap2* in *A. nidulans* leads to production of niduporthin. A) The structure of niduporthin, elucidated by NMR spectrocopy, B) BPC of *A. nidulans* extracts, showing production of niduporthin (EIC @ *m/z* 427.2380 highlighted in blue), and C) BPC of reference strain, which displayed no production of niduporthin.

doi:10.1371/journal.pone.0161199.g004

scenarios we therefore chose to include these two sequences as part of the coding sequence in the purchased gene.

Analogous to the integration and expression of *ccsA* and *ccsC*, *syn2* and *rap2* were transformed in two steps into *A. nidulans* and the metabolite profile of the strain was analyzed by UHPLC-DAD-HRMS. Expression of the two genes resulted in the production of a compound with a mass of 426.2307 Da ($[M+H]^+$ = 427.2380]) corresponding to the formula $C_{28}H_{30}N_2O_2$. Again, it was found that co-expression with the ER was required for detection of a product. The Syn2/Rap2 product was purified and a structure with high resemblance to niduclavin was determined by NMR (Fig 4; S1 Dataset). The Syn2/Rap2 product expressed in *A. nidulans* was named niduporthin. The compound also contained a highly reduced polyketide chain, decalin rings and a nitrogen-containing tetramic acid unit. In conclusion, we wanted to establish whether the two annotated introns that were included in the purchased gene of *syn2* do in fact constitute part of the coding sequence. Therefore, RNA was purified from the *syn2*/*rap2* strain followed by cDNA synthesis. Sequencing of *syn2* cDNA in the regions of the annotated introns confirmed that *A. nidulans* does not splice these two sequences suggesting that they do indeed constitute part of the coding sequence.

The polyketide backbone moiety of niduporthin revealed that the two PKS modules of CcsA and Syn2 perform the same number of iterative elongation steps (seven). However, the methyltransferase (MT) domain of Syn2 attaches only one methyl group to the niduporthin backbone in contrast to the three methyl groups added by the MT domain in CcsA. Furthermore, the polyketide is connected to a tryptophan residue instead of phenylalanine demonstrating different adenylation (A) domain specificities. As for niduclavin, a double bond is found between C-2' and C-3' of niduporthin, which again supports our hypothesis of cross-chemical interactions by endogenous enzyme(s) in *A. nidulans*.

Recently, a product of the Ace1 gene cluster was identified by co-expression of Ace1 with its cognate ER Rap1 in *A. oryzae* [34]. This experiment revealed a highly reduced nonaketide backbone conjugated to a tyrosine moiety, but in contrast to niduclavin, niduporthin, and many other previously described PKS-NRPS products, the compound did not contain a tetramic acid moiety and was instead identified as a linear polyketide where the terminal carboxylic acid group was reduced to an alcohol. It was suggested that this compound was unlikely to be the direct precursor of the final product, but interestingly, the compound was similar to the CcsA/CcsC product described by Oikawa and co-workers [27,35]. In both studies, *A. oryzae* was used as expression hosts. Thus, it seems that *A. oryzae* as well as *A. nidulans* modifies the PKS-NRPS products, albeit in different manners.

Hence, niduporthin is the second purified product of the Ace1 gene cluster, and although deletion of *syn2* was shown to have no significant effect on avirulence or plant host infection,

the final product of the syn2 pathway could still play an accessory function during the infection process. Khaldi *et al*. [36] has proposed that the Ace1/Syn2 gene cluster arose through a partial tandem duplication, and that subsequently, five of the genes, including Syn2, were transferred to an ancestor of *A*. *clavatus* by horizontal gene transfer. Due to structural similarities to the CcsA/CcsC product presented by Fujii *et al*., Cox and co-workers proposed that the Ace1 pathway constitute a "cytochalasan-like" biosynthetic pathway [34]. As shown by Khaldi et al [36] the Syn2-associated part of the cluster is even more closely related to the *ccs* gene cluster, thus suggesting that the product of the syn2 pathway in *M*. *oryzae* is also likely to be a cytochalasan-type of compound. The structure of niduporthin presented in this work, with its close resemblance to the structure of niduclavin, further corroborates this hypothesis.

Investigating NRPS A domain substrate specificity

Exchange of amino acids in fungal polyketide-nonribosomal peptide products have been attempted a number of times by swapping entire NRPS modules [9,16,37]. An alternative and perhaps easier approach when working with these large genes, could be to simply change the A domain specificity by introduction of point mutations. Several models for prediction of the specificity-conferring amino acids of bacterial and fungal NRPS A domains have been published [37–41]. The specificity of bacterial NRPS A domains is well-established; however, the identity of the specificity-conferring amino acids of fungal A domains is much less characterized, which is perhaps due to a more complex mechanism of amino acid selectivity compared to bacterial A domains.

To investigate the specificities of the A domains of CcsA and Syn2, we used the "NRPSpredictor2" web server (http://nrps.informatik.uni-tuebingen.de/) [41,42] to define the signature amino acid residues that are predicted to determine the identity of the amino acid incorporated by the NRPS module. Despite the two A domains sharing only around 48% sequence identity, the ten amino acids predicted to line the active-site binding pocket of CcsA (DMS**E**VG**C**FCK) and Syn2 (DMS**S**VG**G**FCK) vary only in two positions between the enzymes, suggesting that these two residues would explain the observed difference in amino acid specificity. This observation therefore made CcsA and Syn2 an ideal case for studying the quality of the prediction, and of the A domain specificity at the primary structure level. The CcsA A domain uses phenylalanine as substrate while Syn2 uses the larger tryptophan. We attempted to switch the A domain substrate specificities by interchanging the amino acids at these two positions. For each of the enzymes, three strains were constructed; two carrying a single point mutations and one carrying the double mutation.

For CcsA, no significant effect of any point mutations was observed on the production of niduclavin, and we were unable to detect any tryptophan-incorporation in niduclavin in the extract (Fig 5A). In contrast, Syn2-(G3435C) produced only trace amounts of niduporthin, whereas the serine-to-glycine mutation in *syn2* had no effect on niduporthin production. Interestingly, the two mutations combined in Syn2 resulted in a complete loss of product (Fig 5B). In order to facilitate comparisons across strains containing different point mutations, we performed a Southern blot analysis of all these strains to investigate whether the gene copy number could account for the observed differences in the level of product formation. Southern blot analysis indicated that the strain containing Syn2-(G3435C) carried an extra copy of the gene (see S2 Fig). Despite this extra copy, the production of niduporthin was very low compared to unmodified Syn2.

The decrease in product formation observed for Syn2 could perhaps be explained by substitution of smaller amino acids (Ser and Gly) for more bulky amino acids (Glu and Cys), thereby causing a steric clash in the active site, and rendering the enzyme unable to accommodate the more bulky tryptophan substrate. In this scenario, the opposite mutations introduced in CcsA

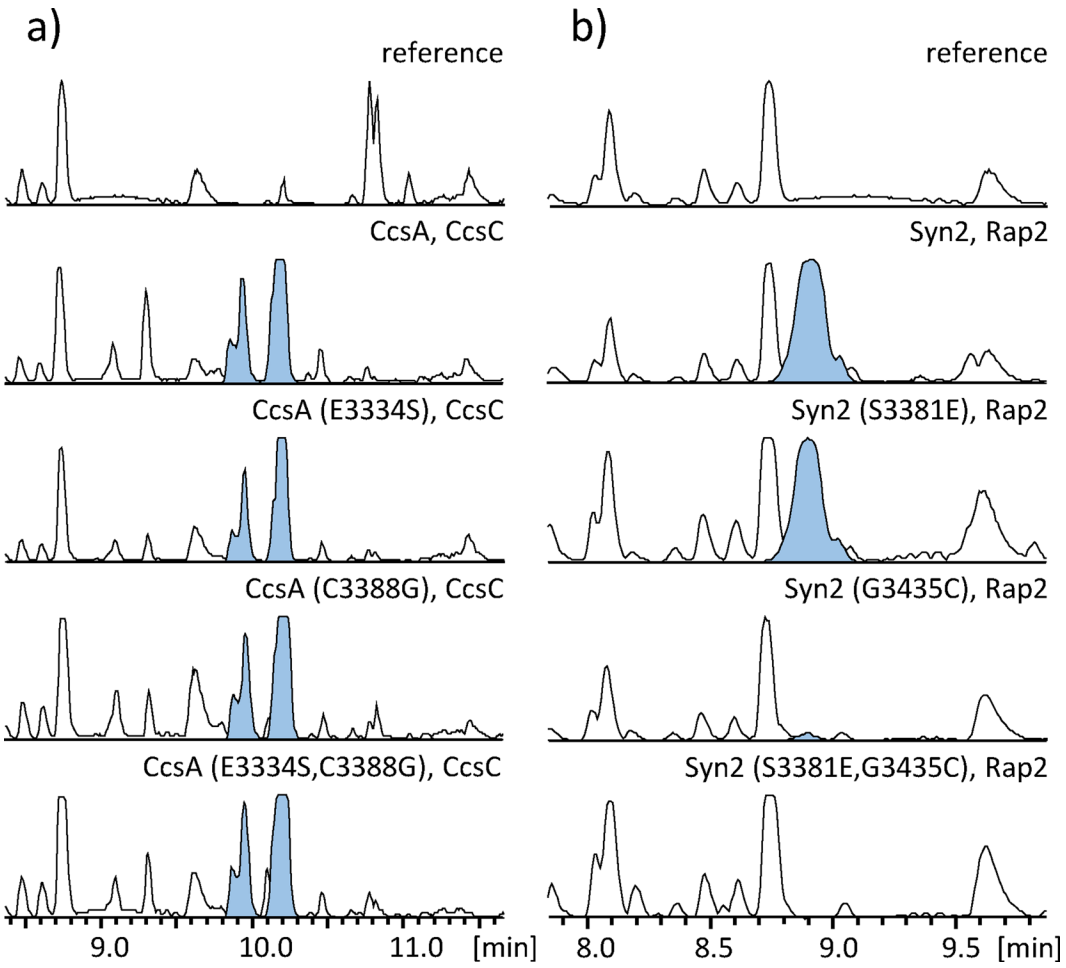


Fig 5. The effects of point mutations in the CcsA- and Syn2 A domains on product formation. The chromatograms show BPCs of *A*. *nidulans* extracts expressing various single- and double mutations in the NRPS A domain as specified above the traces. EICs for niduclavin and niduporthin are highlighted in blue. A) Expression of CcsA containing single- and double point mutations (E3334S and C3388G). B) Expression of Syn2 containing single- and double point mutations (S3381E and G3435C). Only niduporthin-production by Syn2 was affected by introduction of the mutations.

doi:10.1371/journal.pone.0161199.g005

would result in a more spacious binding pocket. It was shown that this change has no effect on niduclavin production but also did not lead to incorporation of tryptophan in place of phenylalanine. The results seem to indicate that the two amino acids are indeed located in the active site binding pocket. However, switching specificities was not achieved, which could suggest that amino acid substrate selection does not solely or directly involve the 10 amino acids predicted by NRPSpredicter2, and that other mechanisms control A domain substrate selectivity. Overall, the rules governing amino acid substrate specificity in fungal NRPSs are likely more complex than for bacterial systems, and it appears that the 10 amino acids predicted for fungal A domains are either not valid, or not sufficient to explain amino acid selectivity.

The CcsA intermodular region shows high tolerance for length and content

One crucial feature for construction of functional chimeric PKS-NRPSs is to understand the transfer mechanisms of polyketide intermediates to the NRPS module. These mechanisms may

include protein-protein interactions *i.e.* compatibility of non-cognate modules along with the mechanism of NRPS substrate selectivity. We considered another mechanism involving linker-mediated polyketide transfer. When examining the domain architecture of CcsA, we observed a region between the acetyl carrier protein (ACP) domain of the PKS module and the condensation (C) domain of the NRPS module containing a stretch of approximately 150 amino acid residues. A BLAST search of the primary sequence of CcsA indicated the presence of an inter-modular region with no homology to other PKS-NRPSs, predicted to display no intrinsic enzymatic activity. We considered an influence of this linker in controlling polyketide chain transfer between the PKS- and NRPS modules, and transfer efficiency would thus be dependent on the length- and amino acid sequence of the linker.

Consequently, we investigated the potential influence of linker composition on products formation by constructing several linker-modified variants of CcsA and Syn2. For linker-swapping, three linkers from different PKS-NRPS homologues were selected (Table 1). The linkers were defined by a combination of alignments and domain predictions, varying in length and sharing no significant sequence similarity to CcsA in the linker region. The linker sequences were PCR amplified and fused to the PKS- and NRPS sequences of *ccsA*, thereby replacing the native linker sequence. The resulting plasmids were transformed into an *A. nidulans ccsC* background strain, and the verified strains were analyzed by UHPLC-DAD-HRMS. The results showed that all linker variants displayed niduclavin production comparable to unmodified CcsA, and the BPCs of all linker-modified variants were comparable to Fig 2B. Thus, no effects of the linker exchange were detected (see S3 Fig). Additionally, four linker-truncated variants of CcsA were constructed (Table 1), however, no effects were observed for any truncations and niduclavin production was retained for all variants. Surprisingly, even the complete removal of the linker, replaced only by a short flexible GSG linker, had no observable effect on product formation. This apparent redundancy of the linker was also tested for Syn2, in which the 163 amino acid linker was replaced with a GSG linker, also showing no significant influence on niduporthin production (see S3 Fig). Finally, a globular protein in the form of red fluorescent protein (RFP) was placed in between the two modules; in one construct situated within the linker, while replacing the linker in another construct. The functionality of the RFP

Table 1. Linker-modified variants of CcsA and Syn2[a].

Strain	Linker modification[b]	Linker length (amino acid)
CcsA WT	None	150
CcsA-CAC	Linker swap–*A. nidulans* AN8412	105
CcsA-CEC	Linker swap–*A. clavatus* ACLA_023380	69
CcsA-CMC	Linker swap–*M. oryzae* CAG28798	163
CcsA-LΔ150	Deletion in the central part of the linker	100
CcsA-LΔ225up	Deletion of the N-terminal end of the linker	75
*CcsAL*Δ225dw	Deletion of the C-terminal end of the linker	75
CcsA-L-GSG	Substitution of the linker for Gly-Ser-Gly	3
Syn2-L-GSG	Substitution of the linker for Gly-Ser-Gly	3
CcsA-RFPlink1	Insertion of RFP in the linker	373
CcsA-RFPlink2	Substitution of the linker for RFP	235

[a] See also S3 Fig.

[b] PKS- and NRPS domains were predicted using the NCBI Conserved Domain Database [43] and the linker was defined between the ACP domain and the condensation domain at positions 2487–2636 (amino acid sequence).

doi:10.1371/journal.pone.0161199.t001

fluorophore was confirmed by fluorescence microscopy (data not shown), and the strains were analyzed by UHPLC-DAD-HRMS. Again, niduclavin production was fully retained. In summary, the results imply that no selection for the linker content exists.

The apparent tolerance for the length of the linker prompted us to test whether niduclavin production could be retained when the two modules were expressed as individual proteins. The dissected PKS- and NRPS modules were co-expressed with the ER CcsC in *A*. *nidulans* and analyzed by UHPLC-DAD-HRMS. The strain did not yield any niduclavin. Compartmentalization of one of the modules could account for the lack of niduclavin production. Therefore, the localizations of the PKS- and NRPS modules were investigated by a C-terminal tagging with RFP and mCitrine, respectively. Both proteins, however, appeared to localize to the cytoplasm (see S4 Fig). This result was surprising since it was previously shown that the activity of the aspyridone PKS-NRPS ApdA could be reconstituted *in vitro* when the two modules were expressed as stand-alone enzymes [8]. However, our results suggest that the primary function of the linker simply is to keep the two modules in close proximity. Furthermore, if PKS-NRPS module-module interactions play a significant role in chain transfer, it seems likely that introduction of a globular protein between the two modules would hamper these interactions. Unexpectedly though, when the linker was replaced by RFP no observable effect on niduclavin production was detected. This could suggest that protein-protein interactions, in fact, do not play an essential role in chain transfer–a conclusion that is in line with the results of Schmidt and co-workers, who showed that the criteria for successful amidation of polyketides are beyond simple ACP-to-C domain interactions [9]. Our results therefore suggest that polyketide substrate recognition by the NRPS module could be the key factor to be considered in construction of chimeric PKS-NRPSs. It also suggests that recombination of PKS-NRPS hybrids *in vivo* and between species may be very flexible, facilitating formation of new spontaneous PKS-NRPSs.

Module swapping of CcsA and Syn2 results in functional chimeric PKS-NRPSs

The linker analysis showed high flexibility and it showed that our definition of PKS and NRPS modules' respective start and ending appeared correct. This allowed us systematically to fuse PKS and NRPS modules from different species to investigate formation of new chimeric hybrid products. We set out to fuse the PKS module of the *A*. *clavatus ccsA* with the NRPS module of the *M*. *oryzae syn2*. To ensure functionality and to provide information on the optimal site for linkage of the two heterologous modules, six variants of the *ccsA-syn2* combination were constructed (see S5 Fig). Among the six *ccsA-syn2* variants, one was joined in the center of the ACP domain to form a hybrid *ccsA-syn2* ACP domain, while another was joined immediately downstream of the ketoreductase (KR) domain. The latter construct was designed to eliminate effects of protein-protein interactions, a strategy also applied by Schmidt and co-workers [9]. All variants were expressed in *A*. *nidulans* along with the ER *ccsC*, and the resulting strains were analyzed by UHPLC-DAD-HRMS. Strikingly, all six variants produced a compound of the mass 454.2620 Da ($[M+H]^+$ = 455.2693) corresponding to an elementary composition of $C_{30}H_{34}N_2O_2$. Indeed, this corresponded exactly to a compound with the polyketide moiety of niduclavin and the tryptophan residue found in niduporthin. Similar to niduclavin and niduporthin, several isomeric compounds were detected in the extracts. The structure of this chimeric compound (named niduchimaeralin A) could be tentatively identified based on tandem MS analysis (see S1 Dataset), since it had a very similar fragmentation pattern to that of niduclavin, including detections of major fragment ions at *m*/z 203 and *m*/*z* 109, strongly indicating that the two compounds have identical decalin-containing polyketide backbones (Fig 6A). The

finding that all six variants are active further supports our previous finding that the linker is highly flexible with regards to length and composition. This also applies to linking heterologous PKS-NRPS fusions, since joining of non-cognate modules lead to the introduction of a non-cognate linker for one of the modules, which in this case lead to a functional enzyme.

To address whether any of the CcsA-Syn2 chimeras were particularly efficient in product formation and to ensure proper comparison of the strains, integration in the intended locus was investigated by Southern blot analysis. This confirmed that all strains had only a single copy of the chimeric gene in the genome and that it had been integrated in the intended integration locus (S2 Fig). This allowed a rough comparison of niduchimaeralin A production for the six CcsA-Syn2 variants based on the relative peak intensities in the base peak chromatograms (S5 Fig). The chimera joined in the center of the linker appeared to display the highest production suggesting that this linkage was the least disruptive of the six. Comparably, production was only slightly lower when the two modules were joined in the middle of the ACP domain, or in the early part of the condensation domain. This is perhaps not surprising since CcsA and Syn2 display conservancy in these regions. On the other hand, linking the two modules immediately downstream of the linker had a substantial negative effect on niduchimaeralin A production. Similarly, joining the two modules downstream of the KR domain led to a loss of product formation (S5 Fig). Effects of protein-protein interactions between non-cognate ACP- and condensation domains were eliminated, and hence, the loss of niduchimaeralin A production must be attributed to unfavorable interactions between the Syn2 ACP domain and the remaining CcsA PKS domains.

The reciprocal chimeric swap where the PKS module of Syn2 was fused to the NRPS module of CcsA was also constructed. The *syn2-ccsA* chimeric gene was transformed into *A*. *nidulans* and co-expressed with the *M*. *oryzae* ER Rap2. Since fusion in the middle of the linker appeared the least disruptive for the CcsA-Syn2 fusion, the Syn2-CcsA chimera was fused in this manner. The product of a functional Syn2-CcsA PKA-NRPS chimera would be expected to contain the PK moiety of Syn2 and the NRP moiety of CcsA. The constructed strain was analyzed by UHPLC-DAD-HRMS, and indeed the presence of at least three isomeric compounds with the expected mass of 387.2198 Da ($[M+H]^+$ = 388.2271) was detected in the extract. Again, this indicates that the decalin ring formation is non-enzymatic. Similar to niduchimeraelin A, tandem MS fragmentation analysis clearly indicated that the structure of the Syn2-CcsA chimeric product, named niduchimaeralin B, contained the polyketide backbone of niduporthin that had likely formed a decalin ring system, fused to a phenylalanine residue of niduclavin (Fig 6B). This was evident from detection of major fragment ions at *m/z* 175 and *m/z* 95 (see S1 Dataset).

The lack of selectivity for the length and content of the linker suggests that protein-protein interactions are not essential for successful amino acid incorporation. This is also in accordance with the findings of Kakule *et al* [9], who constructed a chimeric LovB-EqxS PKS-NRPS where the ACP domain of EqxS was retained. Despite the expected compatibility at the interface between the two modules, the combination did not yield a hybrid polyketide-nonribosomal peptide compound. Consequently, they also hypothesized that C domain substrate selectivity overrules protein-protein interaction.

We speculate that the successful swapping of CcsA and Syn2 modules obtained in this study was achieved, not because of protein structure compatibility, but rather because the products of the two polyketide intermediates are structurally similar. Structural similarity would increase the likelihood that PKS intermediates will be recognized by non-cognate condensation domains and in the case of CcsA and Syn2 will lead to amidation of the polyketide backbone. Our hypothesis is supported by the fact that Syn2 and CcsA do not share overwhelming sequence identity (52%), and yet the polyketide intermediates vary only with two methylations

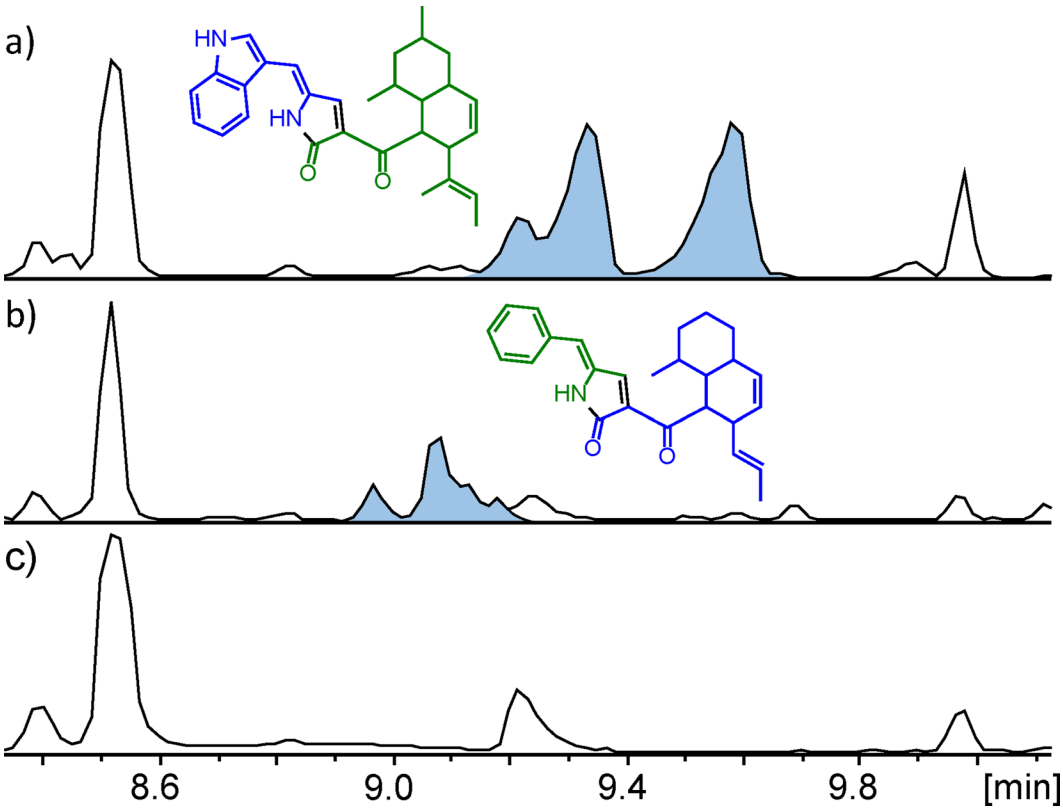


Fig 6. Analysis of chimeric variants of PKS-NRPSs CcsA and Syn2. The chromatograms show BPCs of *A. nidulans* extracts expressing PKS-NRPS hybrid compounds. EICs of the products are highlighted in blue along with structures of the predicted compounds: A) Expression of chimeric *ccsA-syn2* leads to production of niduchimaeralin A (m/z 455.2693). B) Expression of chimeric *syn2-ccsA* leads to production of niduchimaeralin B (m/z 388.2271). c) Reference strain.

doi:10.1371/journal.pone.0161199.g006

of the backbone. If polyketide chain transfer is indeed determined primarily by substrate recognition by the NRPS C domain, and by proximity provided by the inter-modular linker, recombination of even more distantly related PKS-NRPSs should be achievable given that their respective polyketide intermediates are structurally similar.

Materials and Methods

Strains, genomic DNA and media

A list of all the strains used and produced in this study is provided in S1 Table. *Aspergillus nidulans* strain IBT 29539 (*argB2*, *pyrG89*, *veA1*, *nkuAΔ*)—referred to as NID1—was used for heterologous production of niduclavin and niduporthin. *A. clavatus* genomic DNA was obtained from strain IBT 12364 (NRRL 1) and was extracted using the FastDNA™ SPIN Kit for Soil DNA extraction (MP Biomedicals, USA). Coding sequence of *M. oryzae* genes *syn2* and *rap2* were purchased from GenScript USA. *Escherichia coli* strain DH5α was used for plasmid propagation.

Aspergillus solid and liquid minimal medium (MM) and transformation medium (TM) was supplemented when necessary and according to strain genotypes with 4 mM L-arginine, 10 mM uridine, 10 mM uracil, and 1.3 mg/ml 5-fluoroorotic acid (5-FOA), and was prepared as described by Nødvig *et al.* [44]. *E. coli* DH5α was cultivated in Luria-Bertani (LB) medium

consisting of 10 g/l tryptone (Bacto), 5 g/l yeast extract (Bacto), and 10 g/l NaCl (pH 7.0). LB medium was supplemented with 100 μg/ml ampicillin. All solvent used was of HPLC grade, and H_2O was purified and deionized by a Millipore system through a 0.22 μm membrane filter (MQ H_2O).

Vector- and strain construction

All primers (Integrated DNA Technology, Belgium) used in this study are listed in S2 Table. All PCR fragments were generated using the PfuX7 polymerase [45]. All vectors were constructed by Uracil-Specific Excision Reagent (USER) fusion of PCR fragments into compatible vectors [46]. Genes encoding PKS-NRPS hybrids and ERs were cloned into plasmids (pU2115) designed for overexpression by integration into specific targeting sites in the *A. nidulans* genome [47]. The plasmids contain a *Pac*I/*Nt.Bbv*CI USER cassette, the constitutive promoter P*gpdA*, the T*trpC* terminator, *A. nidulans* gene targeting sequences and *A. fumigatus pyrG* flanked by direct repeats for selection and counter selection in *A. nidulans*. For plasmid propagation in *E. coli* the plasmids also contained the *E. coli* ampicillin resistance gene and the origin of replication. The plasmid for deletion of *asqJ* was constructed by introduction of the up- and downstream sequences on each side of the *pyrG* marker. All plasmids were purified using the GenElute^{TM} Plasmid Miniprep Kit (Sigma-Aldrich), and subsequently verified by restriction analysis. For all strains not producing any metabolites, the sequence of the transformed genes were confirmed by sequencing (StarSEQ, Germany) to exclude simple coding errors as reason for not functioning. RNA was purified using the RNeasy Plus Mini Kit from Qiagen, and cDNA was prepared using the Maxima H Minus First Strand cDNA Synthesis Kit from Thermo Scientific^{TM}. All vectors were linearized with *Swa*I (New England Biolabs) prior to transformation according to manufacturer's instructions. *A. nidulans* protoplastation, transformation and rigorous strain validation was performed as described by Nødvig *et al.* [44]. For successive integration of genes, strains were plated on MM supplemented with uridine, uracil, and 5-FOA for counter selection of the *pyrG* marker. The Southern blot protocol is provided in S1 Protocol and the fluorescence microscopy protocol is provided in S2 Protocol.

Chemical analysis

Strains of *A. nidulans* were cultivated at 37°C for 6 days on solid MM with the necessary supplements. Plug extractions were performed as described in Smedsgaard, 1997 [48]. The samples were analyzed on a maXis 3G orthogonal acceleration quadrupole time-of-flight mass spectrometer (Bruker Daltonics) equipped with an electrospray ionization (ESI) source and connected to an Ultimate 3000 UHPLC system (Dionex), equipped with a Kinetex 2.6 μm C18, 100mm x 2.1 mm column (Phenomenex). The method applied was described by Holm *et al.* [49].

Tandem MS experiments were done on an Agilent Infinity 1290 UHPLC system (Agilent Technologies, Santa Clara, CA, USA) equipped with a diode array detector. Separation was obtained on an Agilent Poroshell 120 phenyl-hexyl column (2.1 × 250 mm, 2.7 μm) with a linear gradient consisting of water (A) and acetonitrile (B) both buffered with 20 mM formic acid, starting at 10% B and increased to 100% in 15 min where it was held for 2 min, returned to 10% in 0.1 min and remaining for 3 min (0.35 mL/min, 60°C). MS detection was performed in positive mode on an Agilent 6545 QTOF MS equipped with Agilent Dual Jet Stream electrospray ion source with a drying gas temperature of 250°C, gas flow of 8 L/min, sheath gas temperature of 300°C and flow of 12 L/min. Capillary voltage was set to 4000 V and nozzle voltage to 500 V. Mass spectra were recorded at 10, 20 and 40 eV as centroid data for *m*/*z* 85–1700 in MS mode and *m*/*z* 30–1700 in MS/MS mode, with an acquisition rate of 10 spectra/s. For MS^3,

the fragmentor voltage was increased from 120 V to 200 V, and the desired m/z's (175 and 203) were selected for auto MS/MS. Lock mass solution in 70:30 methanol:water was infused in the second sprayer using an extra LC pump at a flow of 15 μL/min using a 1:100 splitter. The solution contained 1 μM tributylamine (Sigma-Aldrich) and 10 μM Hexakis(2,2,3,3-tetrafluoropropoxy)phosphazene (Apollo Scientific Ltd., Cheshire, UK) as lock masses. The $[M + H]^+$ ions (*m*/*z* 186.2216 and 922.0098 respectively) of both compounds was used. Descriptions of the MS/MS fragmentation patterns of niduclavin, niduporthin, and niduchimaeralin A and B are provided in S1 Text, and for MS/MS data see S1 Dataset.

Purification of Metabolites

For large-scale extracts, strains were cultivated on 6 x 500 mL semi-liquid MM (0.2% agar) at 37°C for 7 days. Extractions were done by separating the mycelium from the media and extracting two times with ethyl acetate (EtOAc); first for one hour with sonication, and second for 12 hrs without sonication. The combined EtOAc phases were dried using a rotary evaporator.

Niduclavin: The extract from the large scale extraction consisting of 0.12 g was adsorbed onto Diol material and dried before packing on a 10 g (~16 mL) SNAP column (Biotage, Uppsala, Sweden) with Diol material. The extract was then fractionated on an Isolera One flash purification system (Biotage) using seven steps of heptane-dichloromethane (DCM)-EtOAc-methanol (MeOH). Fractions were automatically collected one CV at a time. The DCM fractions were subjected to further purification on a semi-preparative HPLC, a Waters 600 Controller with a 996 photodiode array detector (Waters, Milford, MA, USA). This was achieved using a Luna II C18 column (250 × 10 mm, 5 μm, Phenomenex) and 50:50% ACN/H_2O isocratic elution for 5 minutes before increasing to 100% ACN in 15 minutes. The flowrate used was 5 mL/min and 50 ppm TFA of HPLC grade was added to ACN and MQ H_2O. HRMS analysis of the pure compound gave a mass-to-charge ratio of 416.2584, corresponding to a molecular formula of $C_{28}H_{33}NO_2$ (DBE = 13) (calculated for 416.2584, Δ 0 ppm). The yellow amorphous solid displayed UV absorbance at 242 nm and 373 nm (H_2O/MeCN).

Niduporthin: The crude extract (1.2 g) was adsorbed onto C18 material and dried, followed by packing on a 50 g (~66 mL) SNAP column (Biotage) with C18 material. Fractionation was done on an Isolera One flash purification system (Biotage) using a linear MeOH/H_2O-gradient from 0 to 100% MeOH over 32 column volumes. Collection was done automatically using the UV signals at 254 nm and 400 nm with a threshold of 20 mAU. The fractions were analysed by UHPLC-DAD-QTOFMS and the ones containing the desired compound were purified further using the same semi-preparative HPLC system (Waters) and column (Phenomenex) as for niduclavin. The method used was an ACN/H_2O (50 ppm TFA) gradient starting at 65% ACN, increasing to 88% ACN over 8 minutes. Isocratic elution at 88% ACN was done for 8 minutes followed by increasing to 100% ACN over 4 minutes. Analysis by HRMS gave a mass-to-charge of 427.2386, corresponding to a molecular formula of $C_{28}H_{30}N_2O_2$ (DBE = 15) (calculated for 427.2385, Δ -0.2 ppm). The pure compound was a dark red amorphous solid with UV absorption at 228 nm, 270 nm, 286 nm, and 455 nm.

The total yield of niduclavin was 1.1 mg whereas the yield of niduporthin was 23.5 mg. However, it must be noted that niduclavin was purified from a strain carrying the *ccsA* hybrid gene in a different expression site, and we assume a similar yield had niduclavin been purified from the strain carrying *ccsA* in the same expression site as *syn2*.

NMR

All spectra were recorded on a Bruker Avance 800 MHz spectrometer located at the Danish Instrument Centre for NMR Spectroscopy of Biological Macromolecules at Carlsberg

Laboratory. Spectra were acquired using standard pulse sequences. The deuterated solvent was DMSO-d_6 and signals were referenced by solvent signals for DMSO-d_6 at δ_H = 2.50 ppm and δ_C = 39.5 ppm. The NMR data was processed in MestReNova V.10.0.2–15465. Chemical shifts are reported in ppm (δ) and scalar couplings are reported in hertz (Hz). The sizes of the *J* coupling constants in the tables are the experimentally measured values from the 1D ^{1}H and DQF-COSY spectra. There are minor variations in the measurements, which may be explained by the uncertainty of *J* and the spectral digital resolution. Descriptions of NMR structural elucidations of niduclavin and niduporthin are provided in S2 Text and NMR data are provided in S1 Dataset.

Supporting Information

S1 Dataset. Tandem MS and NMR structural elucidation data.
(XLSX)

S1 Fig. Proposed mechanism for [4+2]-cycloaddition of niduclavin. A) Formation of the cytochalasin core structure through a [4+2]-cycloadition as proposed by Qiao et al. [1]. B) Proposed mechanism for the [4+2]-cycloaddition-mediated formation of niduclavin.
(DOCX)

S2 Fig. Southern blot for strains carrying various PKS-NRPS variants. A) The probe hybridizes to the *pyrG* marker and the downstream region. A band of 3.1 kb indicates correct integration of the gene. B) Southern blot of PKS-NRPS variants. Lane1: NID3 reference strain carrying a copy of *pyrG* in the *nkuA* locus, lane 2: *ccsA* WT, lane 3: *syn2* WT, lane 4–9: *ccsA-syn2* chimeric genes, lane 10: *syn2-ccsA* chimeric gene, lane 11–16: various adenylation domain point mutations of ccsA. Unspecific binding (UB) of the probe is observable for samples of high DNA concentration. Since extra copies would be integrated ectopically it is unlikely that two extra bands of identical size for several independent transformants would be seen. Additionally, in the case of extra copies, the intensities of the bands are expected to be equal. Therefore, it was concluded that the bands seen for samples with high DNA concentration represent unspecific binding of the probe. As shown, the *syn2-(GC)* strain carries two copies of the transformed gene. See also Supplemental Experimental Procedures.
(DOCX)

S3 Fig. Base peak chromatograms of A. nidulans extracts expressing various linker modified variants. Base peak chromatograms of A) *ccsA* and B) *syn2*. Highlighted areas represent EICs for A) niduclavin (*m/z* 416.2584), and B) niduporthin (*m/z* 427.2380).
(DOCX)

S4 Fig. Fluorescence tagging of individually expressed CcsA PKS- and NRPS modules. The RFP-tagged PKS module and the mCitrine-tagged NRPS module both appear to be localized to the cytoplasm. C-terminal RFP-tagging of the PKS module and C-terminal mCitrine-tagging of the NRPS module indicate cytoplasmic localization for both modules. N-terminal tagging of the modules revealed the same localization as seen for the C-terminal tagging (results not shown). Scale bar 10 μm. See also Supplemental Experimental Procedures.
(DOCX)

S5 Fig. Constructed fusions between CcsA and Syn2 PKS- and NRPS modules. A) Schematic illustration of the fusions between CccA and Syn2 PKS- and NRPS modules. Arrows indicate the point of fusion. B) Base peak chromatograms of A. nidulans extracts expressing CcsA-Syn2 chimeric PKS-NRPSs. Niduchimaeralin A elutes as several isomeric structures and

are highlighted in blue (EIC @ m/z 455.2693).
(DOCX)

S1 Protocol. Southern blot.
(DOCX)

S2 Protocol. Microscopy.
(DOCX)

S1 Table. List of fungal strains.
(DOCX)

S2 Table. List of primers.
(DOCX)

S1 Text. Fragmentation patterns of niduclavin, niduporthin, and niduchimaeralin A and B.
(DOCX)

S2 Text. NMR structural elucidations of niduclavin and niduporthin.
(DOCX)

Acknowledgments

We thank Kasper Enemark-Rasmussen and Charlotte Held Gotfredsen (Department of Chemistry, Technical University of Denmark) for assistance on running NMR experiments. We also thank Christopher Phippen for valuable discussions during the structural elucidations.

Author Contributions

Conceptualization: MLN UHM MRA JBH TOL.

Formal analysis: MLN TI JBH TOL.

Funding acquisition: TOL.

Investigation: MLN TI LMP.

Writing - original draft: MLN TI.

Writing - review & editing: MLN TI LMP UHM MRA JBH TOL.

References

1. Hertweck C. The Biosynthetic Logic of Polyketide Diversity. Angew Chemie Int Ed. 2009; 48: 4688–4716. doi: 10.1002/anie.200806121
2. Chooi Y-H, Tang Y. Navigating the fungal polyketide chemical space: from genes to molecules. J Org Chem. 2012; 77: 9933–53. doi: 10.1021/jo301592k PMID: 22938194
3. Fisch KM. Biosynthesis of natural products by microbial iterative hybrid PKS–NRPS. RSC Adv. 2013; 3: 18228. doi: 10.1039/c3ra42661k
4. Boettger D, Hertweck C. Molecular diversity sculpted by fungal PKS-NRPS hybrids. Chembiochem. 2013; 14: 28–42. doi: 10.1002/cbic.201200624 PMID: 23225733
5. Heneghan MN, Yakasai AA, Williams K, Kadir KA, Wasil Z, Bakeer W, et al. The programming role of trans-acting enoyl reductases during the biosynthesis of highly reduced fungal polyketides. Chem Sci. 2011; 2: 972. doi: 10.1039/c1sc00023c
6. Heneghan MN, Yakasai A a, Halo LM, Song Z, Bailey AM, Simpson TJ, et al. First heterologous reconstruction of a complete functional fungal biosynthetic multigene cluster. Chembiochem. 2010; 11: 1508–12. doi: 10.1002/cbic.201000259 PMID: 20575135

7. Kakule TB, Sardar D, Lin Z, Schmidt EW. Two related pyrrolidinedione synthetase loci in Fusarium heterosporum ATCC 74349 produce divergent metabolites. ACS Chem Biol. 2013; doi: 10.1021/cb400159f
8. Xu W, Cai X, Jung ME, Tang Y. Analysis of intact and dissected fungal polyketide synthase-nonribosomal peptide synthetase in vitro and in Saccharomyces cerevisiae. J Am Chem Soc. 2010; 5: 13604–13607.
9. Kakule TB, Lin Z, Schmidt EW. Combinatorialization of fungal polyketide synthase – peptide synthetase hybrid proteins. J Am Chem Soc. 2014; 136: 17882–17890. doi: 10.1021/ja511087p PMID: 25436464
10. Seidler NW, Jona I, Vegh M, Martonosi a. Cyclopiazonic acid is a specific inhibitor of the Ca2+-ATPase of sarcoplasmic reticulum. J Biol Chem. 1989; 264: 17816–17823. PMID: 2530215
11. Sondergaard TE, Hansen FT, Purup S, Nielsen AK, Bonefeld-Jørgensen EC, Giese H, et al. Fusarin C acts like an estrogenic agonist and stimulates breast cancer cells in vitro. Toxicol Lett. 2011; 205: 116–121. doi: 10.1016/j.toxlet.2011.05.1029 PMID: 21683775
12. Udagawa T, Yuan J, Panigrahy D, Chang Y, Shah J, Amato RJD. Cytochalasin E, an Epoxide Containing Aspergillus -Derived Fungal Metabolite, Inhibits Angiogenesis and Tumor Growth 1. J Pharmacol Exp Ther. 2000; 294: 421–427. PMID: 10900214
13. Martinez-Luis S, Cherigo L, Arnold E, Spadafora C, Gerwick WH, Cubilla-Rios L. Antiparasitic and anticancer constituents of the endophytic fungus Aspergillus sp. strain F1544. Nat Prod Commun. 2012; 7: 165–168. PMID: 22474943
14. Newman DJ, Cragg GM. Natural Products as Sources of New Drugs from 1981 to 2014. J Nat Prod. 2016; 79: 629–661. doi: 10.1021/acs.jnatprod.5b01055 PMID: 26852623
15. Hertweck C. Decoding and reprogramming complex polyketide assembly lines: prospects for synthetic biology. Trends Biochem Sci. Elsevier Ltd; 2015; 40: 189–199. doi: 10.1016/j.tibs.2015.02.001 PMID: 25757401
16. Fisch KM, Bakeer W, Yakasai AA, Song Z, Pedrick J, Wasil Z, et al. Rational domain swaps decipher programming in fungal highly reducing polyketide synthases and resurrect an extinct metabolite. J Am Chem Soc. 2011; 133: 16635–41. doi: 10.1021/ja206914q PMID: 21899331
17. Ostermeier BM, Benkovic SJ. Evolution of Protein Function by Domain Swapping. Adv Protein Chem. 2001; 55: 29–77.
18. Wriggers W, Chakravarty S, Jennings PA. Control of protein functional dynamics by peptide linkers. Biopolymers. 2005; 80: 736–746. doi: 10.1002/bip.20291 PMID: 15880774
19. Scherlach K, Boettger D, Remme N, Hertweck C. The chemistry and biology of cytochalasans. Nat Prod Rep. 2010; 27: 869–86. doi: 10.1039/b903913a PMID: 20411198
20. Qiao K, Chooi YH, Tang Y. Identification and engineering of the cytochalasin gene cluster from Aspergillus clavatus NRRL 1. Metab Eng. Elsevier; 2011; 13: 723–732. doi: 10.1016/j.ymben.2011.09.008
21. Suzuki S, Hosoe T, Nozawa K, Kawai KI, Yaguchi T, Udagawa SI. Antifungal substances against pathogenic fungi, talaroconvolutins, from Talaromyces convolutus. J Nat Prod. 2000; 63: 768–772. doi: 10.1021/np990371x PMID: 10869198
22. Yang Y-L, Lu C-P, Chen M-Y, Chen K-Y, Wu Y-C, Wu S-H. Cytotoxic Polyketides Containing Tetramic Acid Moieties Isolated from the Fungus Myceliophthora Thermophila: Elucidation of the Relationship between Cytotoxicity and Stereoconfiguration. Chem—A Eur J. 2007; 13: 6985–6991. doi: 10.1002/chem.200700038
23. Kontnik R, Clardy J. Codinaeopsin, an antimalarial fungal polyketide. Org Lett. 2008; 10: 4149–4151. doi: 10.1021/ol801726k PMID: 18698786
24. Campbell CD, Vederas JC. Mini Review: Biosynthesis of lovastatin and related metabolites formed by fungal iterative PKS enzymes. Biopolymers. 2010; 93: 755–763. doi: 10.1002/bip.21428 PMID: 20577995
25. Sims JW, Fillmore JP, Warner DD, Schmidt EW. Equisetin biosynthesis in Fusarium heterosporum. Chem Commun. 2005; 186–8. doi: 10.1039/b413523g
26. Ishikawa N, Tanaka H, Koyama F, Noguchi H, Wang CCC, Hotta K, et al. Non-heme dioxygenase catalyzes atypical oxidations of 6,7-bicyclic systems to form the 6,6-quinolone core of viridicatin-type fungal alkaloids. Angew Chemie Int Ed. 2014; 53: 12880–12884. doi: 10.1002/anie.201407920
27. Fujii R, Minami A, Gomi K, Oikawa H. Biosynthetic assembly of cytochalasin backbone. Tetrahedron Lett. Elsevier Ltd; 2013; 54: 2999–3002. doi: 10.1016/j.tetlet.2013.03.120
28. Kasahara K, Miyamoto T, Fujimoto T, Oguri H, Tokiwano T, Oikawa H, et al. Solanapyrone synthase, a possible Diels-Alderase and iterative type I polyketide synthase encoded in a biosynthetic gene cluster from Alternaria solani. ChemBioChem. 2010; 11: 1245–1252. doi: 10.1002/cbic.201000173 PMID: 20486243

29. Klas K, Tsukamoto S, Sherman DH, Williams RM. Natural Diels-Alderases: Elusive and Irresistable. J Org Chem. 2015; 80: 11672–11685. doi: 10.1021/acs.joc.5b01951 PMID: 26495876

30. Collemare J, Pianfetti M, Houlle A, Morin D, Camborde L, Gagey M-J, et al. Magnaporthe grisea avirulence gene ACE1 belongs to an infection-specific gene cluster involved in secondary metabolism. New Phytol. 2008; 179: 196–208. doi: 10.1111/j.1469-8137.2008.02459.x PMID: 18433432

31. Böhnert HU, Fudal I, Dioh W, Tharreau D, Notteghem J-L, Lebrun M-H. A putative polyketide synthase/peptide synthetase from Magnaporthe grisea signals pathogen attack to resistant rice. Plant Cell. 2004; 16: 2499–2513. doi: 10.1105/tpc.104.022715.1 PMID: 15319478

32. Dean RA, Talbot NJ, Ebbole DJ, Farman ML, Mitchell TK, Orbach MJ, et al. The genome sequence of the rice blast fungus Magnaporthe grisea. Nature. 2005; 434: 980–986. doi: 10.1038/nature03449 PMID: 15846337

33. Xue M, Yang J, Li Z, Hu S, Yao N, Dean R a, et al. Comparative analysis of the genomes of two field isolates of the rice blast fungus Magnaporthe oryzae. PLoS Genet. 2012; 8: e1002869. doi: 10.1371/journal.pgen.1002869 PMID: 22876203

34. Song Z, Bakeer W, Marshall JW, Yakasai AA, Khalid RM, Collemare J, et al. Heterologous expression of the avirulence gene ACE1 from the fungal rice pathogen Magnaporthe oryzae. Chem Sci. Royal Society of Chemistry; 2015; 6: 4837–4845. doi: 10.1039/C4SC03707C

35. Fujii R, Ugai T, Ichinose H, Hatakeyama M, Kosaki T, Gomi K, et al. Reconstitution of biosynthetic machinery of fungal polyketides: unexpected oxidations of biosynthetic intermediates by expression host. Biosci Biotechnol Biochem. 2015; 8451: 1–6. doi: 10.1080/09168451.2015.1104234

36. Khaldi N, Collemare J, Lebrun M-H, Wolfe KH. Evidence for horizontal transfer of a secondary metabolite gene cluster between fungi. Genome Biol. 2008; 9: R18. doi: 10.1186/gb-2008-9-1-r18 PMID: 18218086

37. Boettger D, Bergmann H, Kuehn B, Shelest E, Hertweck C. Evolutionary imprint of catalytic domains in fungal PKS-NRPS hybrids. Chembiochem. 2012; 13: 2363–73. doi: 10.1002/cbic.201200449 PMID: 23023987

38. Stachelhaus T, Mootz HD, Marahiel MA. The specificity-conferring code of adenylation nonribosomal peptide synthetases. Chem Biol. 1999; 6: 493–505. PMID: 10421756

39. Conti E, Stachelhaus T, Marahiel MA, Brick P. Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. EMBO J. 1997; 16: 4174–4183. PMID: 9250661

40. Challis GL, Ravel J, Townsend CA. Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. Chem Biol. 2000; 7: 211–224. PMID: 10712928

41. Röttig M, Medema MH, Blin K, Weber T, Rausch C, Kohlbacher O. NRPSpredictor2—a web server for predicting NRPS adenylation domain specificity. Nucleic Acids Res. 2011; 39: W362–7. doi: 10.1093/nar/gkr323 PMID: 21558170

42. Rausch C, Weber T, Kohlbacher O, Wohlleben W, Huson DH. Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). Nucleic Acids Res. 2005; 33: 5799–5808. doi: 10.1093/nar/gki885 PMID: 16221976

43. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, et al. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res. 2011; 39: D225–D229. doi: 10.1093/nar/gkq1189 PMID: 21109532

44. Nødvig CS, Nielsen JB, Kogle ME, Mortensen UH. A CRISPR-Cas9 system for genetic engineering of filamentous fungi. Yu J-H, editor. One PLoS. 2015; 10: e0133085. doi: 10.1371/journal.pone.0133085

45. Nørholm MHH. A mutant Pfu DNA polymerase designed for advanced uracil-excision DNA engineering. BMC Biotechnol. 2010; 10: 21. doi: 10.1186/1472-6750-10-21 PMID: 20233396

46. Geu-Flores F, Nour-Eldin HH, Nielsen MT, Halkier BA. USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. Nucleic Acids Res. 2007; 35: e55–e55. doi: 10.1093/nar/gkm106 PMID: 17389646

47. Hansen BG, Salomonsen B, Nielsen MT, Nielsen JB, Hansen NB, Nielsen KF, et al. Versatile enzyme expression and characterization system for Aspergillus nidulans, with the Penicillium brevicompactum polyketide synthase gene from the mycophenolic acid gene cluster as a test case. Appl Environ Microbiol. 2011; 77: 3044–51. doi: 10.1128/AEM.01768-10 PMID: 21398493

48. Smedsgaard J. Micro-scale extraction procedure for standardized screening of fungal metabolite production in cultures. J Chromatogr. 1997; 760: 264–270.

49. Holm DK, Petersen LM, Klitgaard A, Knudsen PB, Jarczynska ZD, Nielsen KF, et al. Molecular and chemical characterization of the biosynthesis of the 6-MSA-derived meroterpenoid yanuthone D in Aspergillus niger. Chem Biol. Elsevier Ltd; 2014; 21: 519–529. doi: 10.1016/j.chembiol.2014.01.013 PMID: 24684908

SUPPORTING INFORMATION

Linker Flexibility Facilitates Module Exchange in Fungal Hybrid PKS-NRPS Engineering

Maria Lund Nielsen, Thomas Isbrandt, Lene Maj Petersen¤, Uffe Hasbro Mortensen, Mikael Rørdam Andersen, Jakob Blæsbjerg Hoof, Thomas Ostenfeld Larsen*

Department of Systems Biology, Technical University of Denmark, Søltofts Plads, Kongens Lyngby, Denmark

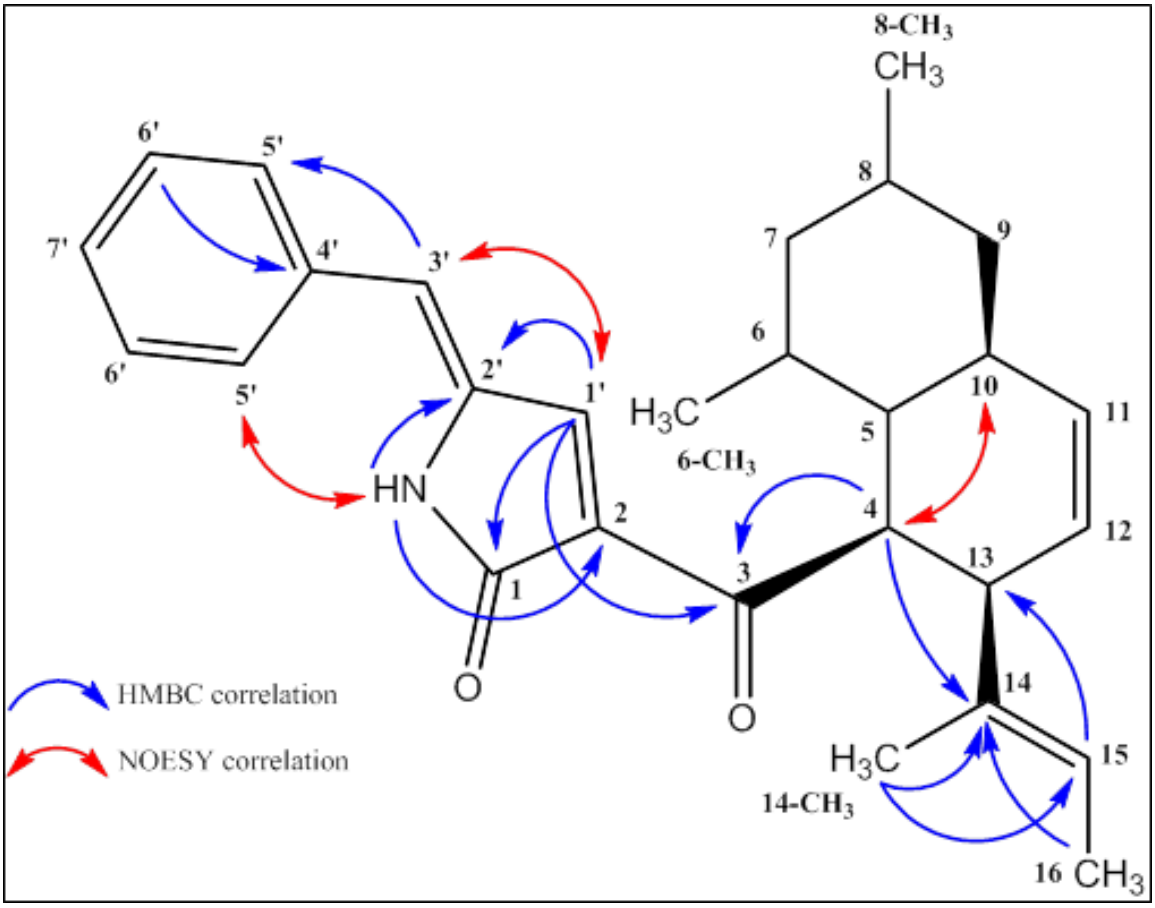
S1. Tandem MS and NMR structural elucidation data

Niduclavin

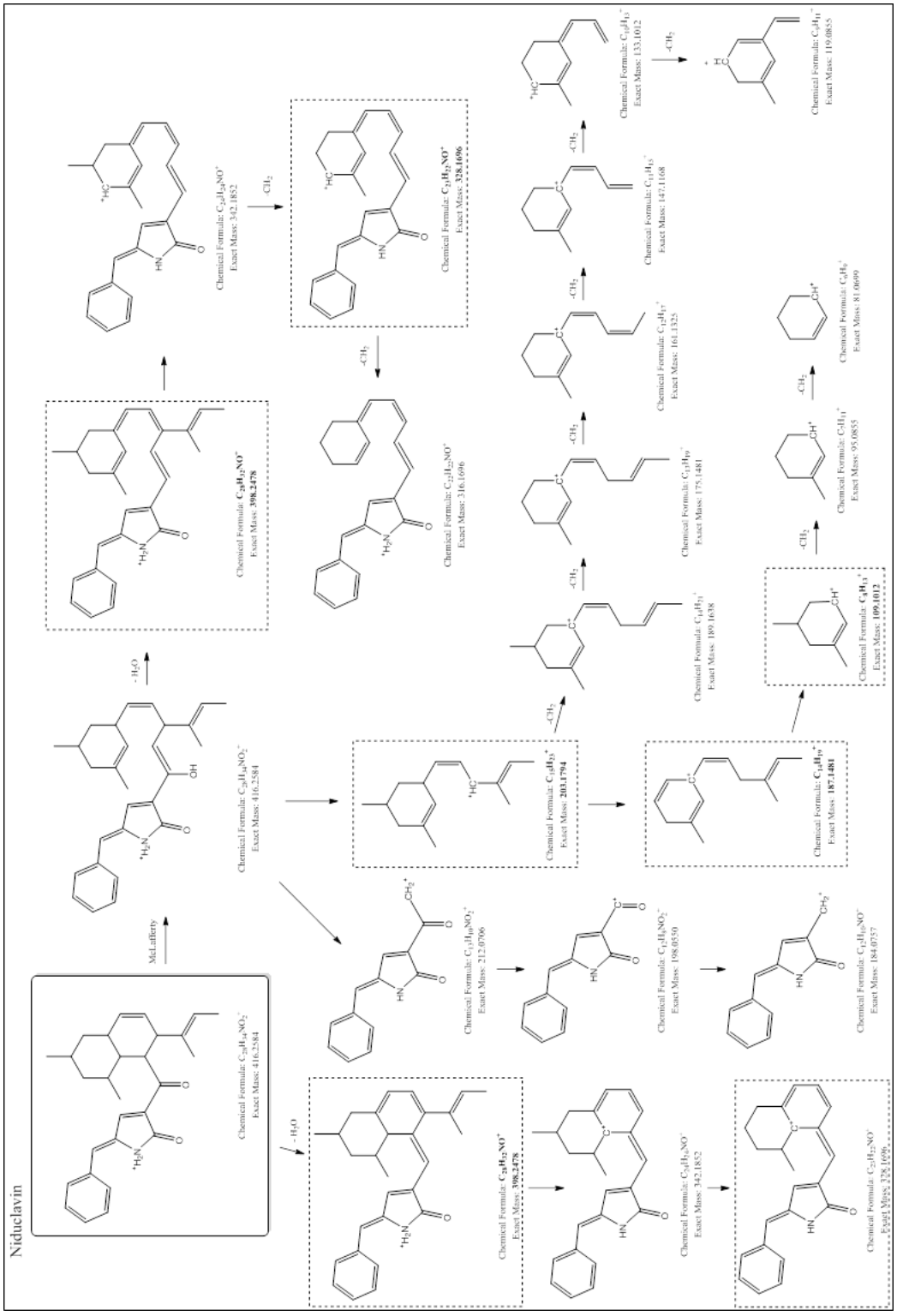
$C_{28}H_{33}NO_2$, m/z = 416.2584 $[M+H]^+$

#	1H	13C (HSQC)	Int. (mult)	J	COSY	HMBC	NOESY
1	-	170.8	C_q	-	-	-	-
2	-	131.8	C_q	-	-	-	-
3	-	197.1	C_q	-	-	-	-
4	3.84	50.7	1H, dd	11.7; 6.7	5, 13	3, 5, 10, 13, 14	10
5	1.75b	39.6	1H, m	-	4, 6, 10	4, 6, 10, 8-CH3	10
6	2.08	28.7	1H, m	-	5, 7a, 7b, 6-CH3	-	7
6-CH3	0.67	14.1	3H, d	7.3	6	5, 6, 7	-
7a	1.49	43	1H, m	-	6, 7b, 8	14-CH3	6
7b	1.16	43	1H, td	12.4; 4.7	6, 7a, 8	6, 8, 6-CH3, 8-CH3	5
8	1.68	27.3	1H, m	-	7, 9, 8-CH3	-	-
8-CH3	0.85	23.6	3H, d	6.35	8	7, 8, 9	-
9a	1.75a	43.2	1H, m	-	8, 9b, 10	-	9b, 10
9b	0.7	43.2	1H, m	12.5	8, 9a, 10	5, 8, 10, 11, 8-CH3	9a
10	2.02	33.9	1H, m	-	5, 9a, 9b, 12	-	4, 5/9a
11	5.58	133.8	1H, m	9.5	12, 13 (W)	5, 9, 10, 13	12
12	5.34	128.6	1H, ddd	9.7; 4.2; 2.6	10, 11, 13	-	11
13	3.31	47.7	1H, m	-	4, 11 (W), 12	-	-
14	-	136.6	C_q	-	-	-	-
14-CH3	1.43	14	3H, d	6.5	15 (W)	4, 14, 15	-
15	5.02	123.4	1H, quartet	6.5	16, 14-CH3 (W)	13, 16, 14-CH3	16
16	1.41	16.8	3H, s	-	15	13, 14, 15	15
1'	7.93	147.1	1H, d	1.8	NH (W)	1, 2, 3, 2', 3'	3'
2'	-	135.9	C_q	-	-	-	-
3'	6.56	120.9	1H, s	-	5'	1', 5'	1', 5'
4'	-	134.9	C_q	-	-	-	-
5'	7.66	131.3	2x 1H, d	7.5	3', 6', 7'	3', 5', 6', 7'	3', 6', NH
6'	7.43	130.1	2x 1H, t	7.7	5', 7'	4', 5', 6', 7'	5'
7'	7.38	130.1	1H, t	7.3	5', 6'	5'	-
NH	10.67	-	NH	-	1' (W)	2, 1', 2'	5'

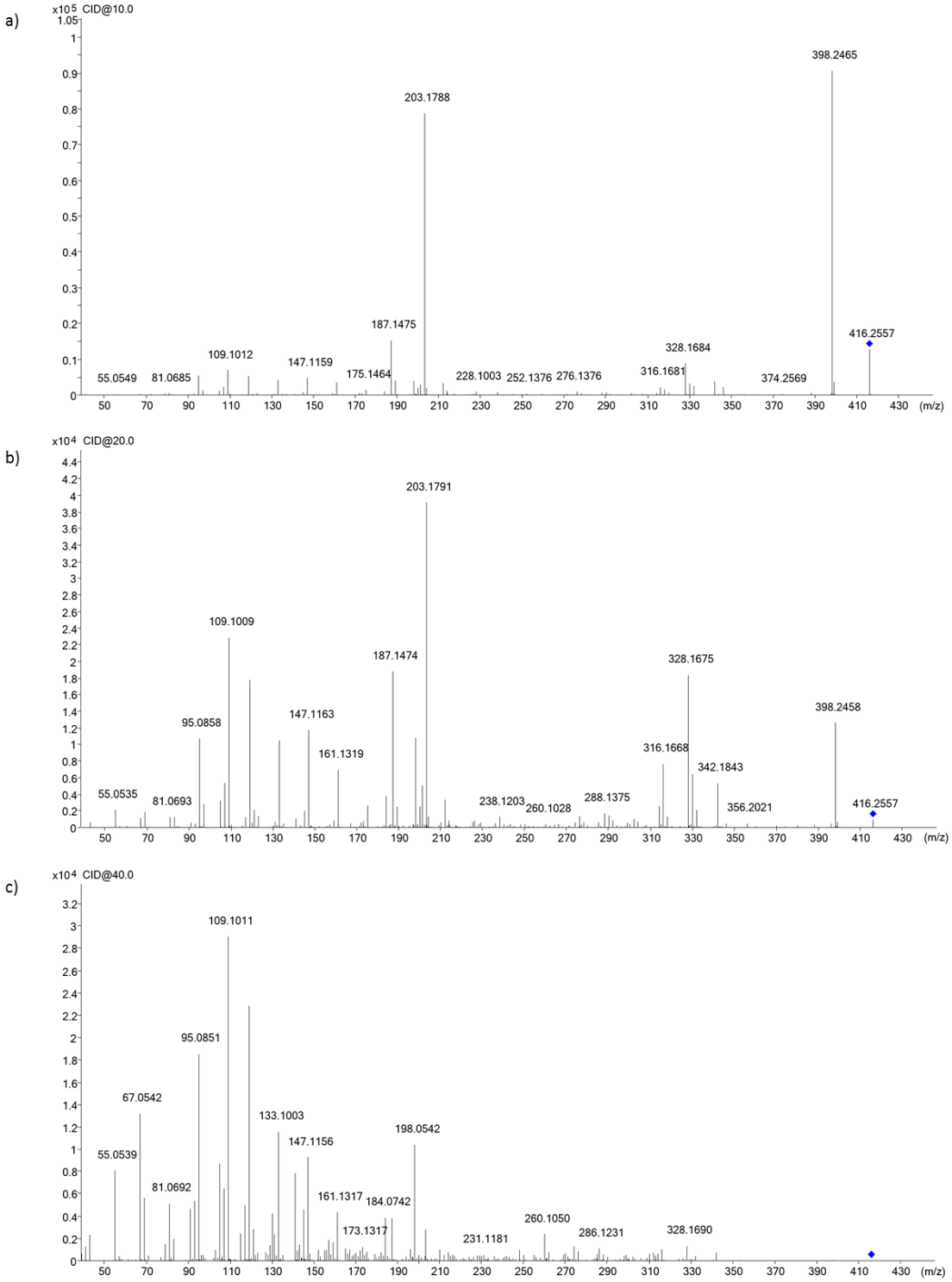
W = W-coupling


8-CH3
CH3
8
9
7
6
10
11
5
12
4
13
3
14
15
16 CH3
H3C
6-CH3
H3C
14-CH3
2
1
HN
O
O
6'
5'
7'
4'
3'
6'
5'
2'
1'
HMBC correlation
NOESY correlation

MS/MS fragmentation - Suggested fragmentation patterns for niduclavin. Fragments corresponding to peaks of major intensities are framed.



MS/MS spectra of niduclavin at different collision energies. a) 10 eV, b) 20 eV, and c) 40 eV.

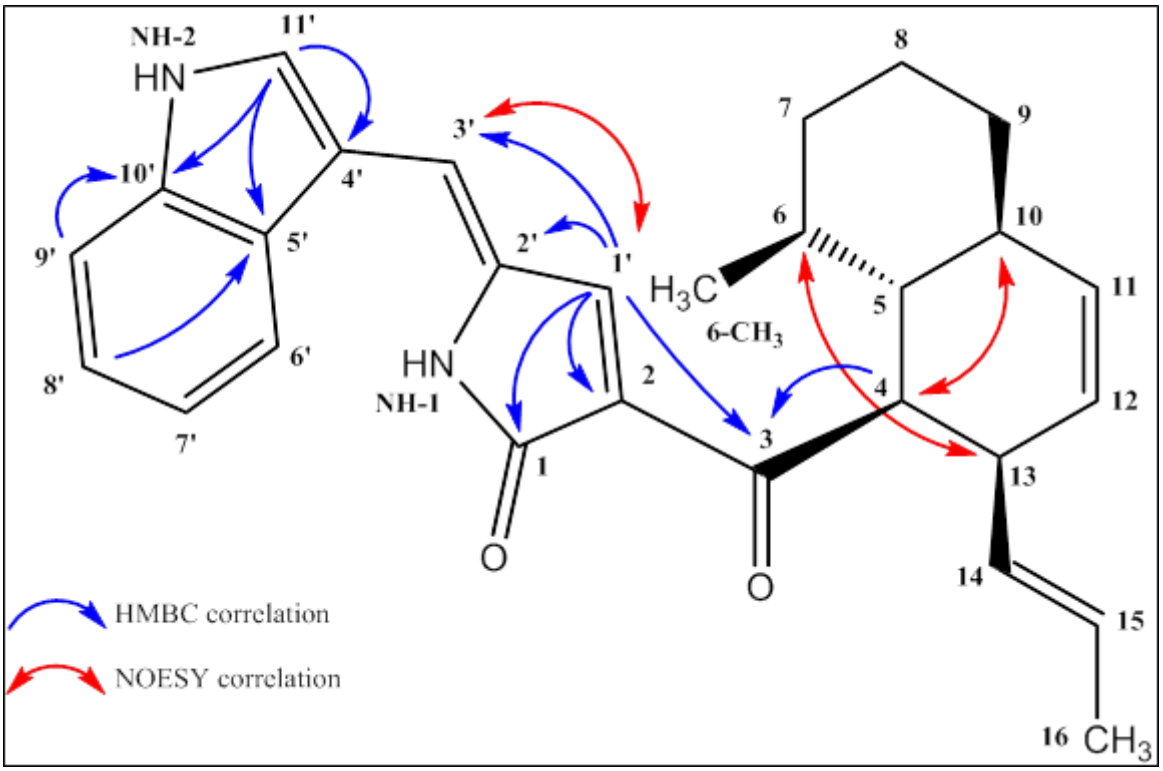


Niduporthin

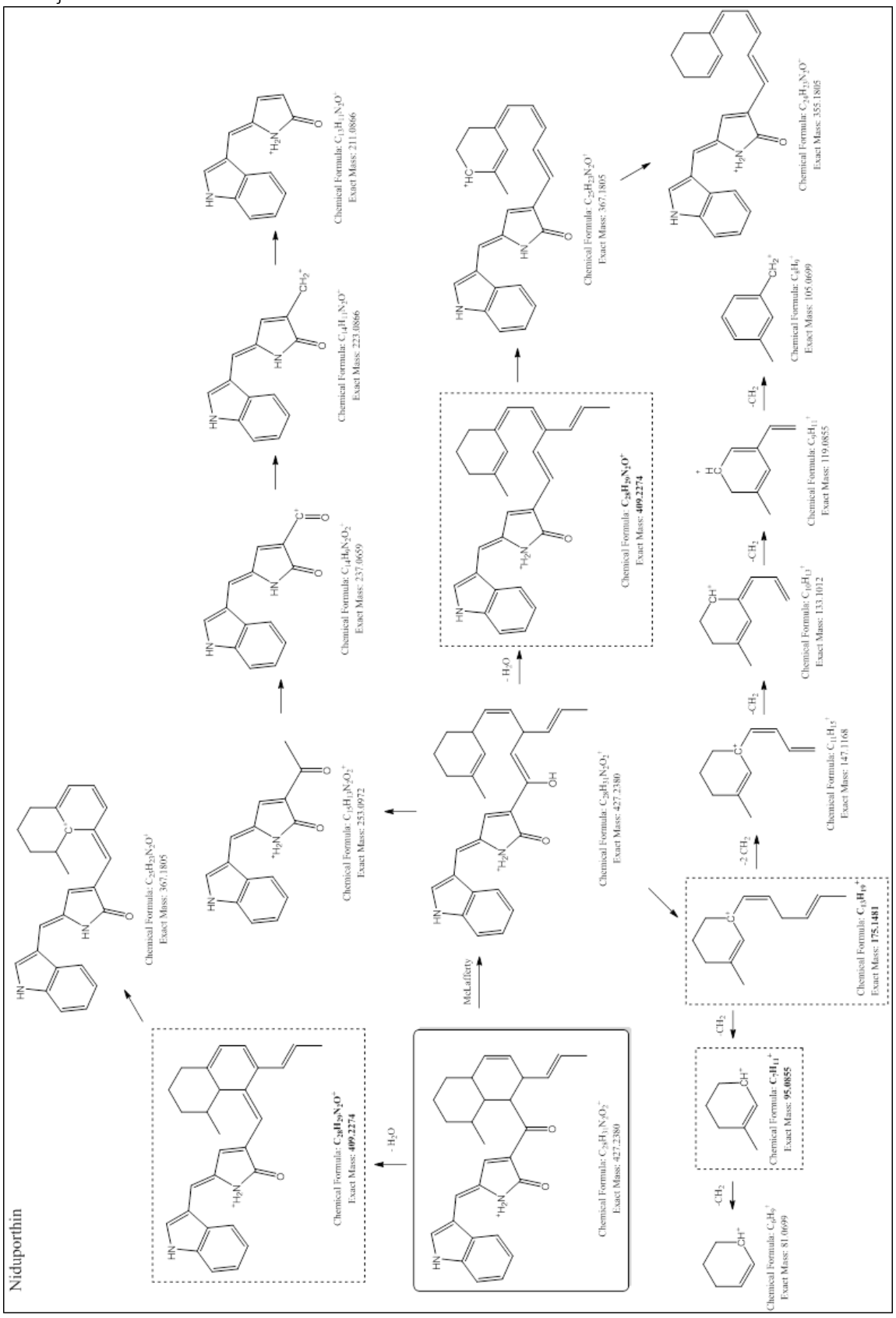
$C_{28}H_{30}N_2O_2$, m/z = 427.2380 $[M+H]^+$

#	1H	13C (HSQC)	Int. (mult)	J	COSY	HMBC	NOESY
1	-	169.2	C_q	-	-	-	-
2	-	128.47	C_q	-	-	-	-
3	-	197.46	C_q	-	-	-	-
4	4.08	50.21	1H, dd	10; 7.5	13, 5	3, 5, 12, 13, 14	6, 10
5	1.45	45.31	1H, m	-	4, 6, 10	4, 6, 9, 10	-
6	1.3	38.84	1H, m	-	5, 7b	-	4, 7a
6-CH_3	0.67	22.7	3H, d	6.6	6	5, 6, 7	7b
7a	1.61	36.45	1H, m	-	7b, 8, 9a (W)	8, 9	6
7b	1.08	36.45	1H, m	-	6, 7a, 8	5, 6, 8, 9, 6-CH_3	6-CH3
8a	1.7	26.01	1H, m	-	7, 8b, 9	-	8b
8b	1.4	26.01	1H, m	-	7, 8a, 9	-	8a
9a	1.84	32.93	1H, m	-	7 (W), 8, 9b, 10	5, 7, 8	9b
9b	1.16	32.93	1H, m	-	8, 9a, 10	5, 8, 10, 11	9a
10	1.76	40.68	1H, m	-	9, 5	5, 9	4
11	5.55	132.68	1H, m	-	-	5, 9, 13	-
12	5.5	128.61	1H, m	-	-	4, 13	-
13	2.99	40.1	1H, m	-	4	12, 14, 15	-
14	5.2	131.93	1H, m	-	13	12, 13, 16	-
15	5.23	124.45	1H, m	-	16	13, 14, 16	-
16	1.48	17.67	3H, d	5.5	15	14, 15	-
1'	7.81	144.06	1H, m	-	-	1, 3, 2', 3'	3'
2'	-	130.49	C_q	-	-	-	-
3'	7.01	114.15	1H, s	-	-	1', 5', 11'	1'
4'	-	110.76	C_q	-	-	-	-
5'	-	127.1	C_q	-	-	-	-
6'	7.8	118.14	1H, m	-	7'	8', 10'	7'
7'	7.18	120.77	1H, m	-	6'	5', 9'	6'
8'	7.22	122.83	1H, m	-	9'	6', 10'	9'
9'	7.45	112.18	1H, m	-	8'	5', 7'	8'
10'	-	135.93	C_q	-	-	-	-
11'	8.38	129.09	1H, d	2.8	NH-2	4', 5', 10'	NH-1, NH-2
NH-1	10.33	-	1H, m	-	-	2, 1', 2', 10'	11'
NH-2	12.15	-	1H, d	2	11'	4', 5', 10', 11'	9', 11'

W = W-coupling

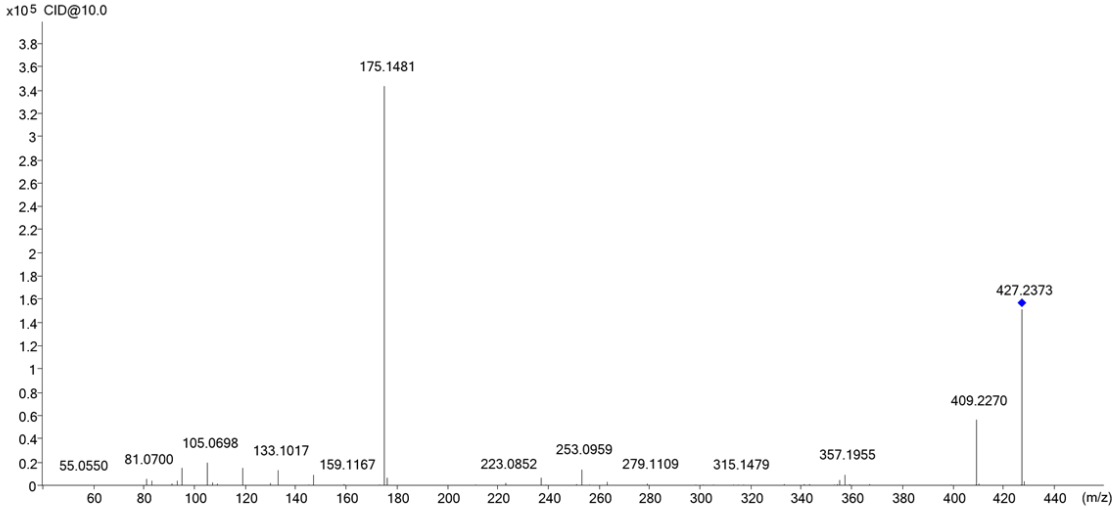

NH-2
HN
11'
10'
9'
8'
7'
6'
5'
4'
3'
2'
1'
HN
NH-1
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16 CH3
H3C
6-CH3
O
O
HMBC correlation
NOESY correlation

MS/MS fragmentation - Suggested fragmentation patterns for niduporthin. Fragments corresponding to peaks of major intensities are framed.

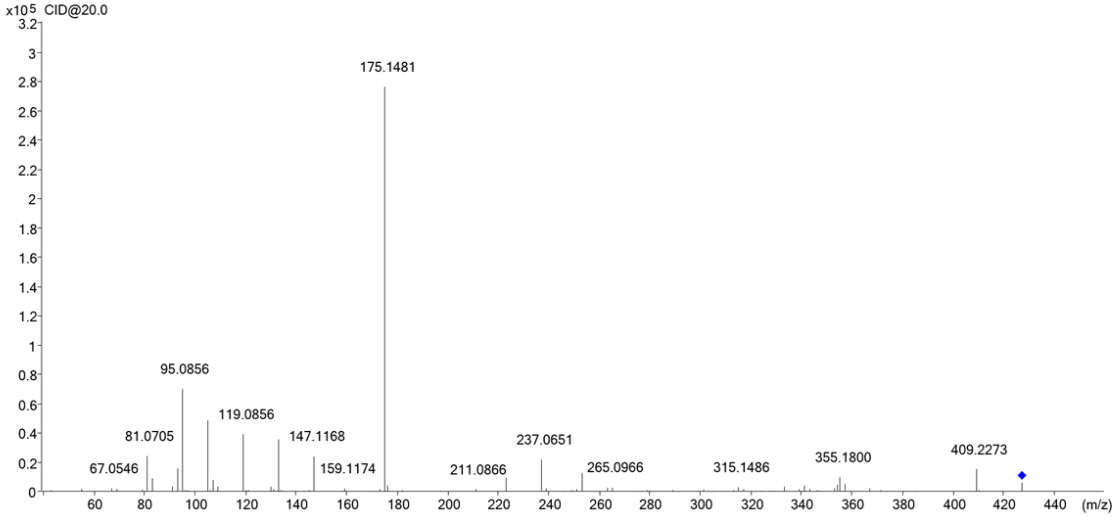


MS/MS spectra of niduporthin. a) 10 eV, b) 20 eV, and c) 40 eV.

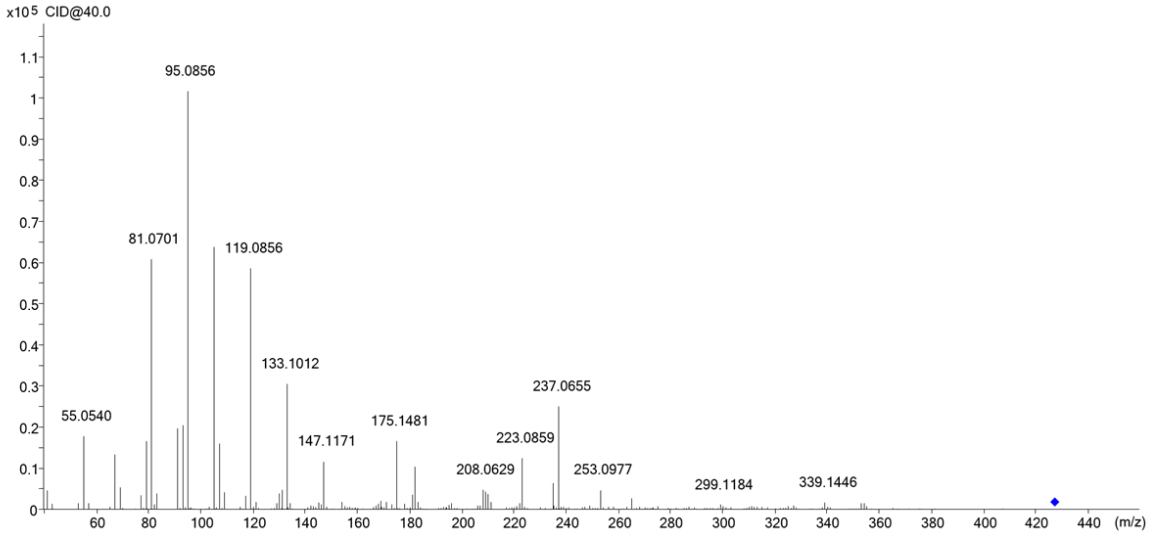
a)



b)

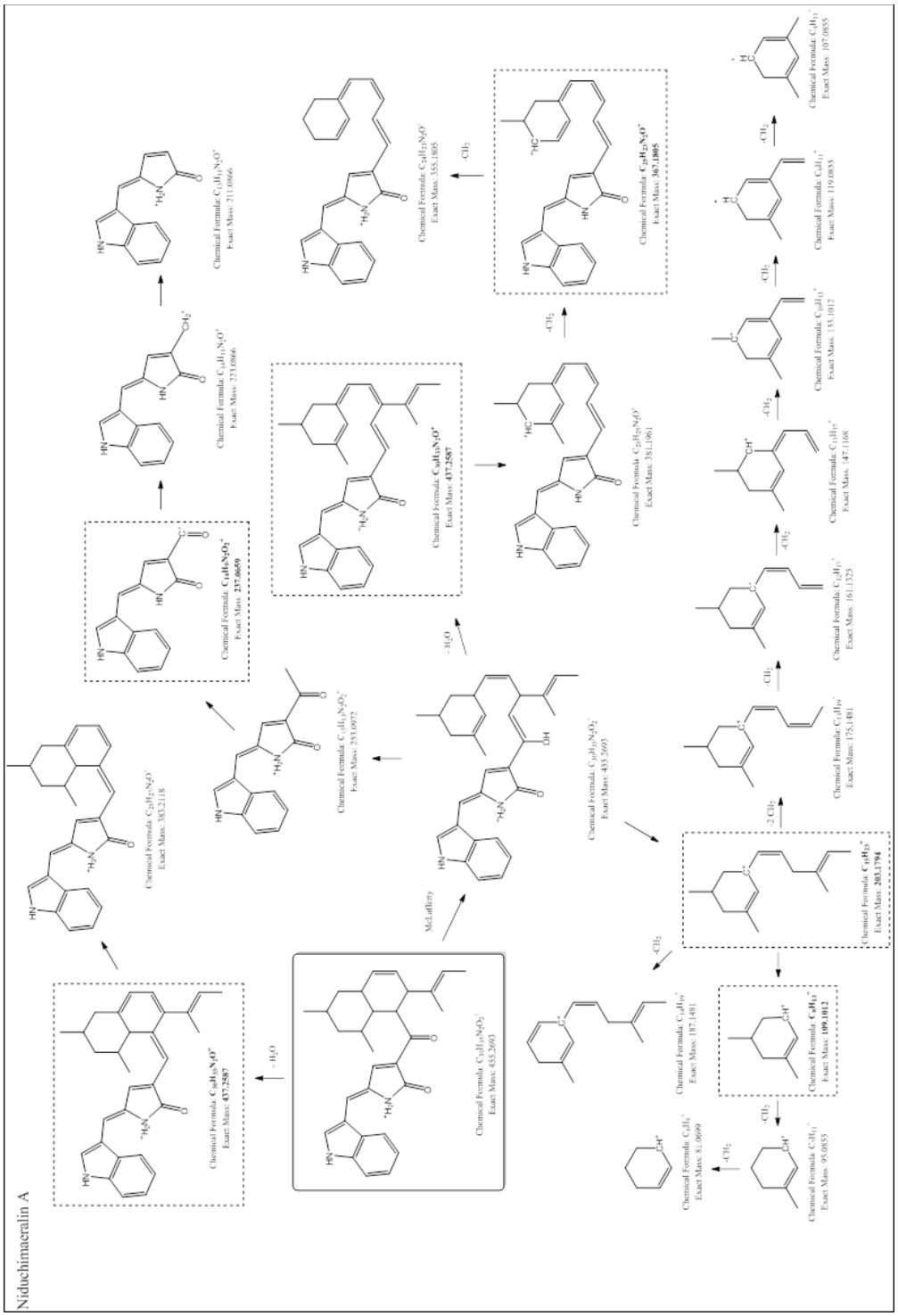


c)

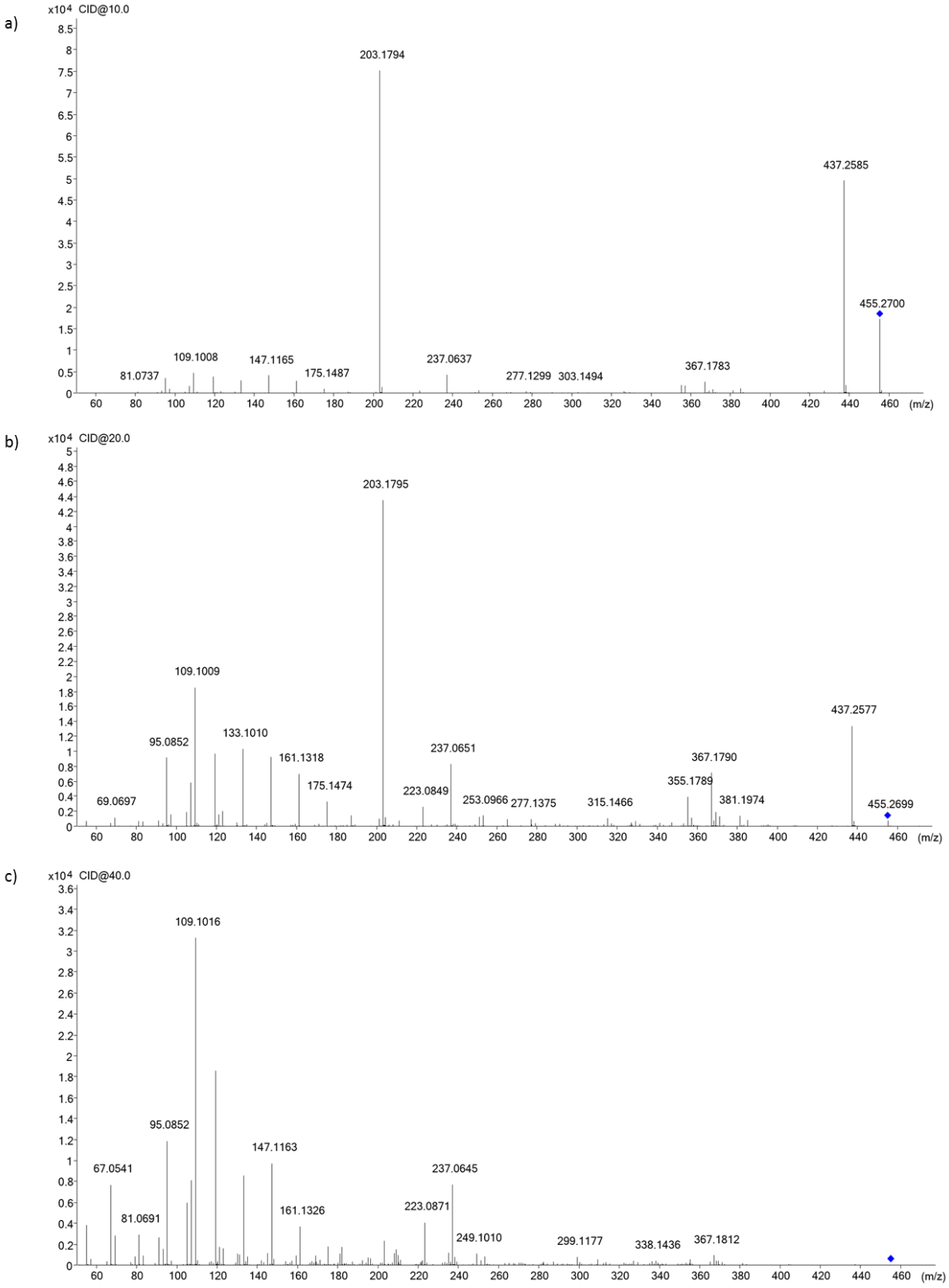


Niduchimaeralin A

MS/MS fragmentation - Suggested fragmentation patterns for niduchimaeralin A. Fragments corresponding to peaks of major intensities are framed.

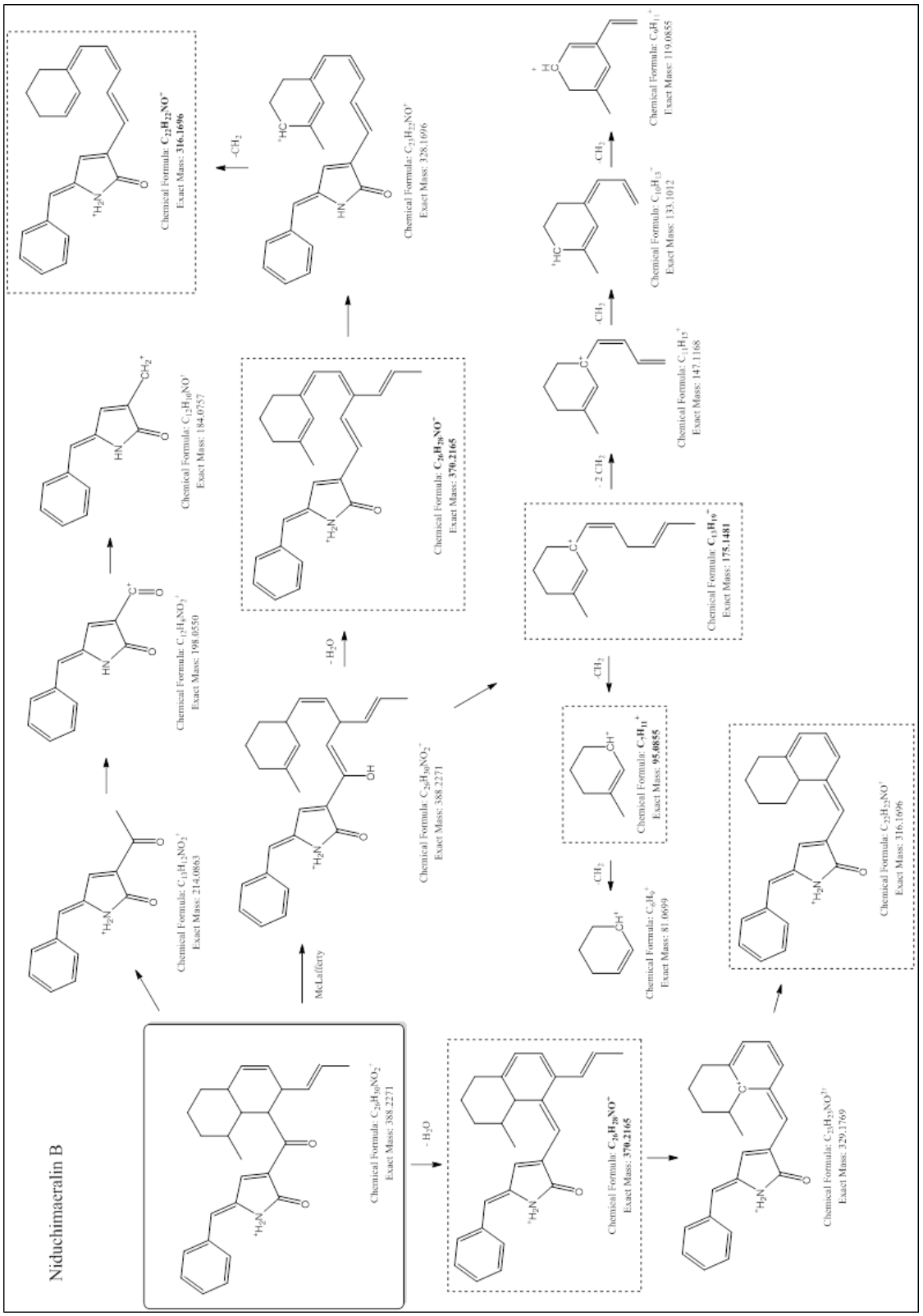


MS/MS spectra of niduchimaeralin A. a) 10 eV, b) 20 eV, and c) 40 eV.

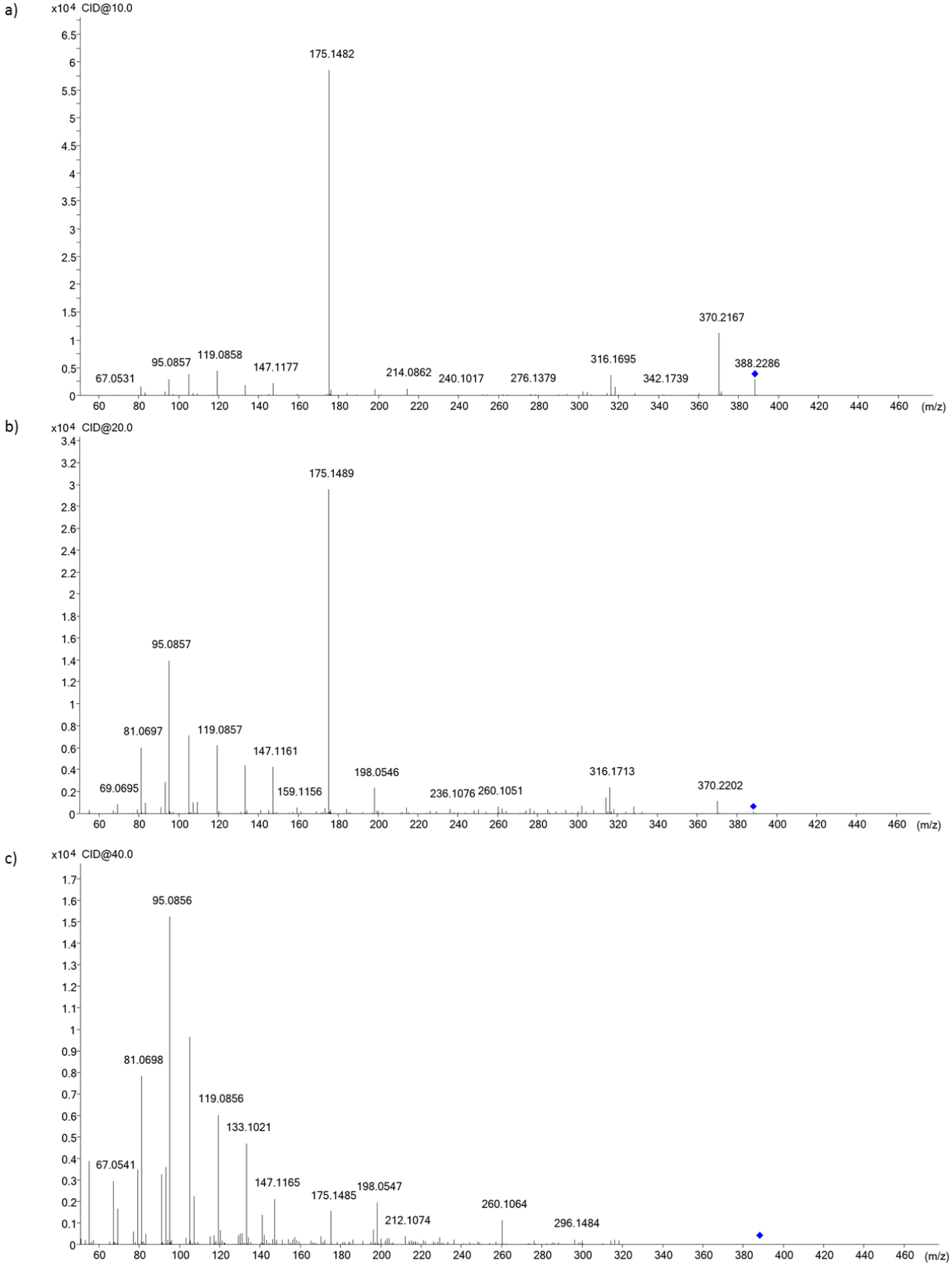


Niduchimaeralin B

MS/MS fragmentation - Suggested fragmentation patterns for niduchimaeralin B. Fragments corresponding to peaks of major intensities are framed.

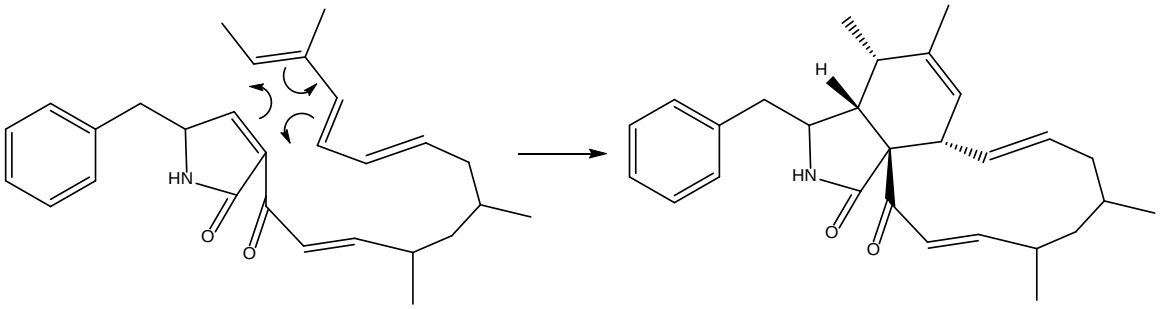


MS/MS spectra of niduchimaeralin B. a) 10 eV, b) 20 eV, and c) 40 eV.

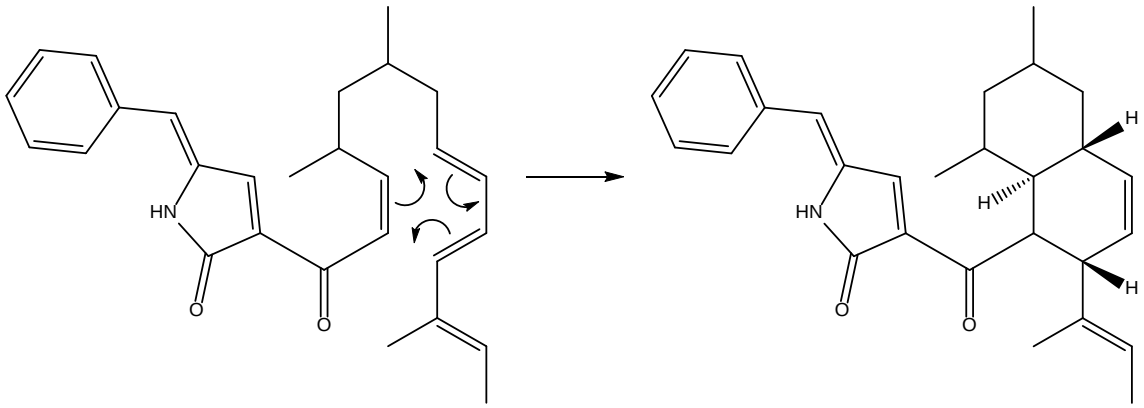


S1 Fig. Proposed mechanism for [4+2]-cycloaddition of niduclavin. A) Formation of the cytochalasin core structure through a [4+2]-cycloadition as proposed by Qiao *et al.* [1]. B) Proposed mechanism for the [4+2]-cycloaddition-mediated formation of niduclavin.

A)



B)



1. Qiao K, Tang Y, Chooi Y-H. Identification and engineering of the cytochalasin gene cluster from Aspergillus clavatus NRRL 1. Metab Eng. Elsevier; 2011;13: 723–732. doi:10.1016/j.ymben.2011.09.008

S1 Protocol. Southern blot

DNA for southern blot was extracted as follows: Spores were harvested from plates of MM supplemented with arginine and grown for 2 days in shake flasks at 37°C. The biomass was filtered and freeze-dried overnight, and the freeze-dried mycelium was grinded in a mortar. For each strain, 10 ml lysis buffer was added consisting of 3.75 ml of buffer A (0.35 M sorbitol, 0.1 M Tris-HCl pH 9, 5 mM EDTA pH 8), 3.75 ml of pre-heated (65°C) buffer B (0.2 M Tris-HCl pH 9, 50 mM EDTA pH 8, 2 M NaCl, 2 % CTAB), 1.5 ml 5 % Sarkosyl, 1 ml 1 % PVP and 100 µl Proteinase K. The samples were vortexed and incubated at 65°C for 30 minutes followed by the addition of 3.35 ml of 5 M potassium acetate, and subsequent incubation on ice for 30 minutes. The lysates were centrifuged for 30 minutes at 5000 g at 4°C and 5 ml of phenol:chloroform:iso-amylalchohol (25:24:1) was added to the supernatant. The samples were centrifuged for 20 minutes at 4°C and the aqueous phase was transferred to new tubes and 1/10 volume of 3 M sodium acetate, and 1 volume of isopropanol was added followed by centrifugation for 30 minutes at 4°C. The pellet was washed with 2 ml 70 % ethanol and finally re-dissolved in 600 µl TE buffer.

For each southern blot sample, 2 µg of DNA was digested with *Pst*I, and the blot was performed as described by Sambrook and Russell [1]. For generation of the probe, a 692 bp DNA fragment was PCR amplified using the primers AFpyrG-F and AFpyrG-R and the Biotin DecaLabel DNA Labelling Kit (Thermo Scientific) was used for incorporation of biotin. Detection was achieved using the Biotin Chromogenic detection kit (ThermoFisher Scientific) using streptavidin conjugated to alkaline phosphatase.

1. Sambrook, J., Russell DW. Molecular cloning a laboratory manual. Cold Spring Habor, NY.: Cold Spring Harbor Laboratory Press; 2001.

S1 Table. List of fungal strains

#	Strain identifier	Genotype
1	nkuAΔ	*argB2, pyrG89, veA1, nkuAΔ*
2	nkuA-trS	*argB2, pyrG89, veA1, nkuA-trS::pyrG*
3	ccsA prepop	*argB2, pyrG89, veA1, nkuAΔ, IS1::PgpdA-ccsA-TtrpC::AFpyrG*
4	syn2 prepop	*argB2, pyrG89, veA1, nkuAΔ, IS5::PgpdA-syn2-TtrpC::AFpyrG*
5	ccsC prepop	*argB2, pyrG89, veA1, nkuAΔ, IS2::PgpdA-ccsC-TtrpC::AFpyrG*
6	ccsC pop	*argB2, pyrG89, veA1, nkuAΔ, IS2::PgpdA-ccsC-TtrpC*
7	rap2 prepop	*argB2, pyrG89, veA1, nkuAΔ, IS2::PgpdA-rap2-TtrpC::AFpyrG*
8	rap2 pop	*argB2, pyrG89, veA1, nkuAΔ, IS2::PgpdA-rap2-TtrpC*
9	ccsA, ccsC prepop	*argB2, pyrG89, veA1, nkuAΔ, IS5::PgpdA-ccsA-TtrpC::AFpyrG, IS2::PgpdA-ccsC-TtrpC*
10	syn2, rap2 prepop	*argB2, pyrG89, veA1, nkuAΔ, IS5::PgpdA-syn2-TtrpC::AFpyrG, IS2::PgpdA-rap2-TtrpC*
11	ccsA-syn2-1, ccsC prepop	*argB2, pyrG89, veA1, nkuAΔ, IS5::PgpdA-CM1-TtrpC::AFpyrG, IS2::PgpdA-ccsC-TtrpC*
12	ccsA-syn2-2, ccsC prepop	*argB2, pyrG89, veA1, nkuAΔ, IS5::PgpdA-CM2-TtrpC::AFpyrG, IS2::PgpdA-ccsC-TtrpC*
13	ccsA-syn2-3, ccsC prepop	*argB2, pyrG89, veA1, nkuAΔ, IS5::PgpdA-CM3-TtrpC::AFpyrG, IS2::PgpdA-ccsC-TtrpC*
14	ccsA-syn2-4, ccsC prepop	*argB2, pyrG89, veA1, nkuAΔ, IS5::PgpdA-CM4-TtrpC::AFpyrG, IS2::PgpdA-ccsC-TtrpC*
15	ccsA-syn2-5, ccsC prepop	*argB2, pyrG89, veA1, nkuAΔ, IS5::PgpdA-CM5-TtrpC::AFpyrG, IS2::PgpdA-ccsC-TtrpC*
16	ccsA-syn2-6, ccsC prepop	*argB2, pyrG89, veA1, nkuAΔ, IS5::PgpdA-CM6-TtrpC::AFpyrG, IS2::PgpdA-ccsC-TtrpC*
17	syn2-ccsA, rap2 prepop	*argB2, pyrG89, veA1, nkuAΔ, IS5::PgpdA-MC6-TtrpC::AFpyrG, IS2::PgpdA-rap2-TtrpC*
18	ΔasqJ, ccsA, ccsC prepop	*argB2, pyrG89, veA1, nkuAΔ, AN9227Δ::AFpyrG, IS1::PgpdA-ccsA-TtrpC, IS2::PgpdA-ccsC-TtrpC*
19	CAC, ccsC prepop	*argB2, pyrG89, veA1, nkuAΔ, IS5::PgpdA-CAC-TtrpC::AFpyrG, IS2::PgpdA-ccsC-TtrpC*
20	CEC, ccsC prepop	*argB2, pyrG89, veA1, nkuAΔ, IS5::PgpdA-CEC-TtrpC::AFpyrG, IS2::PgpdA-ccsC-TtrpC*
21	CMC, ccsC prepop	*argB2, pyrG89, veA1, nkuAΔ, IS5::PgpdA-CMC-TtrpC::AFpyrG, IS2::PgpdA-ccsC-TtrpC*
22	ccsA(LΔ150), ccsC prepop	*argB2, pyrG89, veA1, nkuAΔ, IS5::PgpdA-ccsA(LD150)-TtrpC::AFpyrG, IS2::PgpdA-ccsC-TtrpC*
23	ccsA(LΔ225up), ccsC prepop	*argB2, pyrG89, veA1, nkuAΔ, IS5::PgpdA-ccsA(LD225up)-TtrpC::AFpyrG, IS2::PgpdA-ccsC-TtrpC*
24	ccsA(LΔ225dw), ccsC prepop	*argB2, pyrG89, veA1, nkuAΔ, IS5::PgpdA-ccsA(LD225dw)-TtrpC::AFpyrG, IS2::PgpdA-ccsC-TtrpC*
25	ccsA(L-GSG), ccsC prepop	*argB2, pyrG89, veA1, nkuAΔ, IS5::PgpdA-ccsA(L-GSG)-TtrpC::AFpyrG, IS2::PgpdA-ccsC-TtrpC*
26	syn2(L-GSG), rap2 prepop	*argB2, pyrG89, veA1, nkuAΔ, IS5::PgpdA-syn2(L-GSG)-TtrpC::AFpyrG, IS2::PgpdA-rap2-TtrpC*
27	ccsA-RFPlink1, ccsC prepop	*argB2, pyrG89, veA1, nkuAΔ, IS5::PgpdA-ccsA-RFPlink1-TtrpC::AFpyrG, IS2::PgpdA-ccsC-TtrpC*
28	ccsA-RFPlink2, ccsC prepop	*argB2, pyrG89, veA1, nkuAΔ, IS5::PgpdA-ccsA-RFPlink2-TtrpC::AFpyrG, IS2::PgpdA-ccsC-TtrpC*
29	ccsA(PKS)-RFP prepop	*argB2, pyrG89, veA1, nkuAΔ, IS5::PgpdA-ccsA(PKS)-RFP-TtrpC::AFpyrG*
30	ccsA(NRPS)-mCitrine prepop	*argB2, pyrG89, veA1, nkuAΔ, IS2::PgpdA-ccsA(NRPS)-mCitrine-TtrpC::AFpyrG*

#	Strain notes
1	Permanent *nku* deletion strain (background strain)
2	Transient small repeat in *nkuA* (reference strain for chemical analysis)
3	Oex of *A. clavatus ccsA* (ACLA_078660) from the *ccs* gene cluster
4	Oex of *M. oryzae* PKS-NRPS *syn2* (CAG_28798)
5	Oex of *A. clavatus* enoyl reductase *ccsC* (ACLA_078700) from the *ccs* gene cluster
6	Oex of *A. clavatus* enoyl reductase *ccsC* (ACLA_078700) from the *ccs* gene cluster
7	Oex of *M. oryzae* enoyl reductase *rap2* (MGG_08380)
8	Oex of *M. oryzae* enoyl reductase *rap2* (MGG_08380)
9	Oex of *A. clavatus ccsA* (ACLA_078660) and *ccsC* (ACLA_078700)
10	Oex of *M. oryzae syn2* (CAG_28798) and *rap2* (MGG_08380)
11	Oex of chimeric PKS-NRPS, *ccsA* PKS fused with *syn2* NRPS moiety
12	Oex of chimeric PKS-NRPS, *ccsA* PKS fused with *syn2* NRPS moiety
13	Oex of chimeric PKS-NRPS, *ccsA* PKS fused with *syn2* NRPS moiety
14	Oex of chimeric PKS-NRPS, *ccsA* PKS fused with *syn2* NRPS moiety
15	Oex of chimeric PKS-NRPS, *ccsA* PKS fused with *syn2* NRPS moiety
16	Oex of chimeric PKS-NRPS, *ccsA* PKS fused with *syn2* NRPS moiety
17	Oex of chimeric PKS-NRPS, *syn2* PKS fused with *ccsA* NRPS moiety
18	Deletion of AN9227 in *ccsA, ccsC* background strain
19	Oex of *ccsA* with linker swapped for *A. nidulans apdA* (AN8412), co-expressed with enoyl reductase *ccsC*
20	Oex of *ccsA* with linker swapped for *A. clavatus eqiS* (ACLA_023380) linker, co-expressed with enoyl reductase *ccsC*
21	Oex of *ccsA* with linker swapped for *M. oryzae syn2*, co-expressed with enoyl reductase *ccsC*
22	Oex of *ccsA* with truncated linker, co-expressed with enoyl reductase *ccsC*
23	Oex of *ccsA* with truncated linker, co-expressed with enoyl reductase *ccsC*
24	Oex of *ccsA* with truncated linker, co-expressed with enoyl reductase *ccsC*
25	Oex of *ccsA* without linker (only GSG linker), co-expressed with enoyl reductase *ccsC*
26	Oex of *syn2* without linker (only GSG linker), co-expressed with enoyl reductase *rap2*
27	Oex of *A. clavatus ccsA* with RFP in the linker, co-expressed with enoyl reductase *ccsC*
28	Oex of *A. clavatus ccsA* with RFP as the linker, co-expressed with enoyl reductase *ccsC*
29	Oex of RFP-tagged PKS module from *ccsA*
30	Oex of mCitrine-tagged NRPS module from *ccsA*

S1 Text. Fragmentation patterns of niduclavin, niduporthin, and niduchimaeralin A and B.

Based on high resolution tandem MS experiments it was possible to give tentative structures for the two hybrid products niduchimaeralin A and B, when compared to those for niduclavin and niduporthin. MS^2 was done on all four products to help verify the fragmentation patterns observed. In addition, MS^3 experiments were performed on the ions *m/z* 175 and *m/z* 203, to further verify the fragmentation patterns of these ions (data not shown).

As niduclavin and niduporthin are proposed to be formed via [4+2] cycloadditions, retro-Diels Alder derived fragments were expected. This, however, was not the case. Rather it would seem that a McLafferty rearrangement provided the basis for most fragments as have also been showing for the structurally related chaetoglobosins (Xu et al., 2012).

Both niduclavin and niduporthin show fragments in line with an initial McLafferty rearrangement, followed by a variety of fragmentation pathways. The most predominant fragments for **niduclavin** are water loss, and water loss followed by loss of a butenyl group, both of which would be possible with or without a preceding McLafferty rearrangement. Furthermore, an ion with a mass-to-charge ratio of 203 is expected to correspond to part of the polyketide after a McLafferty rearrangement, with additional fragments matching consecutive CH_2 losses (*m/z* = 189, 175, 161, 147, 133, and 119), also confirmed by MS^3-experiments.

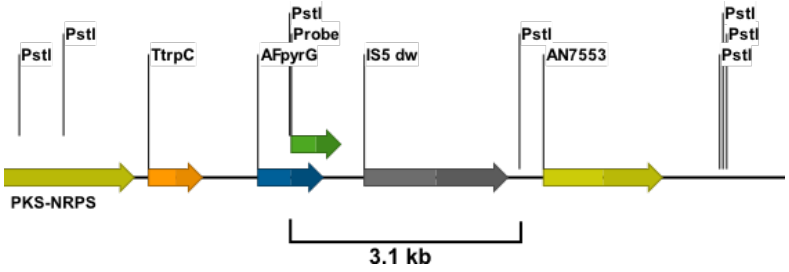
The major fragments in the **niduporthin** spectra are water loss, and a peak corresponding to water loss followed by propene loss, similar to what was observed for niduclavin. Moreover, a fragment with an m/z of 175 matches fragmentation following a McLafferty rearrangement, followed by ion matching succeeding CH_2 losses (*m/z* = 147 (2x CH_2), 133, 119, 105), also confirmed by MS^3-experiments.

Niduchimaeralin A was expected to share PK backbone with niduclavin, and differ by the incorporation of tryptophan instead of phenylalanine. Both compounds share a fragment ion with an *m/z* of 203 as the most predominant peak, and both followed by similar CH_2 losses (*m/z* = 14). Additionally, a single water loss ion (*m/z* = 437), along with additional loss of butenyl is observed (*m/z* = 367).

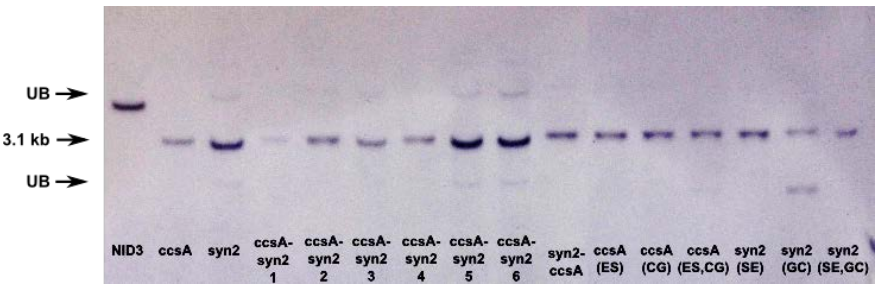
Niduchimaeralin B was expected to be the reverse construction of niduchimaeralin A, i.e. the niduporthin PK backbone with phenylalanine incorporated. When comparing the fragmentation pattern to that of niduporthin similar fragments are observed. The major fragment is a peak with an m/z of 175, followed by fragments matching losses of CH_2. Similarly to niduchimaeralin A water loss is also observed as a major fragment.

S2 Fig. Southern blot for strains carrying various PKS-NRPS variants. A) The probe hybridizes to the *pyrG* marker and the downstream region. A band of 3.1 kb indicates correct integration of the gene. B) Southern blot of PKS-NRPS variants. Lane1: NID3 reference strain carrying a copy of *pyrG* in the *nkuA* locus, lane 2: *ccsA* WT, lane 3: *syn2* WT, lane 4-9: *ccsA-syn2* chimeric genes, lane 10: *syn2-ccsA* chimeric gene, lane 11-16: various adenylation domain point mutations of ccsA. Unspecific binding (UB) of the probe is observable for samples of high DNA concentration. Since extra copies would be integrated ectopically it is unlikely that two extra bands of identical size for several independent transformants would be seen. Additionally, in the case of extra copies, the intensities of the bands are expected to be equal. Therefore, it was concluded that the bands seen for samples with high DNA concentration represent unspecific binding of the probe. As shown, the *syn2-(GC)* strain carries two copies of the transformed gene. See also Supplemental Experimental Procedures.

A)



B)



S2 Protocol. Microscopy

MM agar slides were prepared by covering with 1 ml MM agar (with necessary supplements), then inoculated with spores and incubated at 30°C in petri dishes overnight. A cooled Evolution QEi monochrome digital camera (Media Cybernetics Inc.) mounted on a Nikon Eclipse E1000 microscope (Nikon) captured live-cell images using a Plan-Fluor x100, 1.30 numerical aperture objective lens. The illumination source was a 103-watt mercury arc lamp (Osram). The fluorophores RFP and mCitrine were visualized using a band pass RFP (EX545/30, EM620/60; Nikon) and YFP filter (EX500/20, EM535/30; Nikon), respectively. Exposure time for images was 500 msec. Red and yellow colors were added to the corresponding fluorescence signals using image processing in ImageJ.

S2 Table. List of Primers

Name	Sequence	Purpose
ACLAccsA-eOex-1FU_2	GGGTTTAAUATGGGGTCATTTCAGAACTCCTC	ccsA into pU2115-5
ACLAccsA-eOex-2FU	AGAGAGGAAUGGACTGGAATGGCATCACAAAC	ccsA into pU2115-5
ACLAccsA-eOex-2RU	ATTCCTCTCUCTCTCGGGAACCGCTTCAC	ccsA into pU2115-5
ACLAccsA-eOex-3FU	AAGGATTUGGCCGCAGAGGAGAAGCA	ccsA into pU2115-5
ACLAccsA-eOex-3RU	AAATCCTUAGCCACAACAATGTTCCGTGA	ccsA into pU2115-5
ccsA-Pac-RU	GGTCTTAAUTGCTGTGTCCCAATCAGACGT	ccsA into pU2115-5
ACLA-ccsC-FU	AGAGCGAUATGACCGTACCAACCACTATCCG	ccsC into pU2111-2 (p60)
ACLA-ccsC-RU	TCTGCGAUTTACATGCCGATGCTCACAACC	ccsC into pU2111-2 (p60)
SYN2-Pac-FU	GGGTTTAAUATGGAAGCCCAGAATGAGCC	syn2 into pU2115-5
SYN2-RU	ACCGAGACTCUTGCCAGAGCGTTCGGCT	syn2 into pU2115-5
SYN2-2FU	AGAGTCTCGGUTGTCAAAAGCCTTCAGGACGA	syn2 into pU2115-5
SYN2-2RU	AAACAAAAGUCCGAACCTCATCTGCTCGTT	syn2 into pU2115-5
SYN2-3FU	ACTTTTGTTUGGCGGTCTTCCTGGAGC	syn2 into pU2115-5
SYN2-3RU	ATGAGTTCCUTGGGCAGCAGCTTCTGAG	syn2 into pU2115-5
SYN2-4FU	AGGAACTCAUACCCAAATTCGATGAGAAGG	syn2 into pU2115-5
SYN2-4RU	ACCAAAGUGGCGGGCCATTGTGG	syn2 into pU2115-5
SYN2-5FU	ACTTTGGUGCATCGCATCGACGACAT	syn2 into pU2115-5
SYN2-Pac-5RU	GGTCTTAAUTCAATGCCAAGCCTCTCCC	syn2 into pU2115-5
RAP2-Pac-FU	GGGTTTAAUATGTATATTCCTTCGGCGAGGA	rap2 into pU2115-2
RAP2-Pac-RU	GGTCTTAAUCTATTCAAGTGATACAACCAGCTTCTG	rap2 into pU2115-2
ACLAccsA-eOex-1FU_2	GGGTTTAAUATGGGGTCATTTCAGAACTCCTC	CM chimeras into pU2115-5
ACLAccsA-eOex-2FU	AGAGAGGAAUGGACTGGAATGGCATCACAAAC	CM chimeras into pU2115-5
ACLAccsA-eOex-2RU	ATTCCTCTCUCTCTCGGGAACCGCTTCAC	CM chimeras into pU2115-5
SYN2-Pac-5RU	GGTCTTAAUTCAATGCCAAGCCTCTCCC	CM chimeras into pU2115-5
CM1-mid-RU	AGCGTTGGGAUCCAGGTTCGGG	CM chimeras into pU2115-5
CM1-end-FU	ATCCCAACGCUACTGGCCTCAGCAAGC	CM chimeras into pU2115-5
CM2-mid-RU	ACCGGTTTGGUCTTTGTAATGATTTGGCTGGCA	CM chimeras into pU2115-5
CM2-end-FU	ACCAAACCGGUGGAGATCACCAAGAGCGGG	CM chimeras into pU2115-5
CM3-mid-RU	ATGGCACCAAUATGCACAGC	CM chimeras into pU2115-5
CM3-end-FU	ATTGGTGCCAUCTTTGGCAACGGCTACGTCA	CM chimeras into pU2115-5
CM4-mid-RU	ACCAGCGAGUCAATTCCCAA	CM chimeras into pU2115-5
CM4-end-FU	ACTCGCTGGUGGCCGTTGACATCCGATCCT	CM chimeras into pU2115-5
CM5-mid-RU	AGGGCCTCGUGACGCTGTCCGACGACTTGAA	CM chimeras into pU2115-5
CM5-end-FU	ACGAGGCCCUCCGCACC	CM chimeras into pU2115-5
CM6-mid-RU	ATCTTGCTAUCCTTTGGTGCAGCCT	CM chimeras into pU2115-5
CM6-end-FU	ATAGCAAGAUCTCCAGCCCCAAATCC	CM chimeras into pU2115-5
CM7-mid-RU	AGCTTCGGGGUCAACTCCTTGGG	CM chimeras into pU2115-5
CM7-end-FU	ACCCCGAAGCUGCAGGAGGCTTTGC	CM chimeras into pU2115-5
SYN2-Pac-FU	GGGTTTAAUATGGAAGCCCAGAATGAGCC	MC6 chimera into pU2115-5
SYN2-2RU	AAACAAAAGUCCGAACCTCATCTGCTCGTT	MC6 chimera into pU2115-5
SYN2-3FU	ACTTTTGTTUGGCGGTCTTCCTGGAGC	MC6 chimera into pU2115-5
ccsA-Pac-RU	GGTCTTAAUTGCTGTGTCCCAATCAGACGT	MC6 chimera into pU2115-5

SYN2-6RU	ATATTGGTCAUCAGCTGGTATCAACACTGAC	MC6 chimera into pU2115-5
ccsA-10FU	ATGACCAATAUCCAACAGCACTTGAGACTC	MC6 chimera into pU2115-5
ccsA-7RU	ATTGGGAUCCAGGTTCGGG	Linker swap/modification in ccsA
ccsA_NRPS-2FU	AAGAAGAGTGUTCCTATGGCGTTT	Linker swap/modification in ccsA
SYN2_linker-FU	ATCCCAAUGCTACTGGCCTCAGCAAGC	Linker swap/modification in ccsA
SYN2_linker-RU	ACACTCTTCTUGACCTCGAAGCCGGCGCTGGC	Linker swap/modification in ccsA
EqiS_linker-FU	ATCCCAAUGCGCTAACAAACACCAAGATTG	Linker swap/modification in ccsA
EqiS_linker-RU	ACACTCTTCTUGACCTCCTTAGCTGCCTGTGAGACTTTAATATT	Linker swap/modification in ccsA
ApdA_linker-FU	ATCCCAAUCAGCTGGAGAAGCAAGACACG	Linker swap/modification in ccsA
ApdA_linker-RU	ACACTCTTCTUGACCTCGGAGGCAGGCTCAGCG	Linker swap/modification in ccsA
CNC_linker-FU	ATCCCAAUTCCCCTGCGACAGTGTCG	Linker swap/modification in ccsA
CNC_linker-RU	ACACTCTTCTUGACCTCCGCGGGAGAATCGGGTT	Linker swap/modification in ccsA
ccsA-8RU	ATGCTTCTCAUTGGTCTTGATCATGTTCTGCAC	Linker swap/modification in ccsA
ccsA-3FU	ATGAGAAGCAUCTCACCGATCAGGAACCCG	Linker swap/modification in ccsA
ccsA-9RU	AAGTGCTGUGTTGGGATCCAGGTTCGGG	Linker swap/modification in ccsA
ccsA-4FU	ACAGCACTUGAGACTCCATCAAAG	Linker swap/modification in ccsA
ccsA-10RU	ACCTCTGGAUATTGGCTATCCTTTGGTG	Linker swap/modification in ccsA
ccsA-5FU	ATCCAGAGGUCAAGAAGAGTGTTCCTATGG	Linker swap/modification in ccsA
ccsA-11RU	ACCTCGCCUGAGCCGTTGGGATCCAGGTTCGGG	Linker swap/modification in ccsA
ccsA-6FU	AGGCGAGGUCAAGAAGAGTGTTCCTATGG	Linker swap/modification in ccsA
asqJup-FU	GGGTTTAAUGTTTTGGTAGAGAAGAATGGATGG	asqJ deletion in A. nidulans
asqJup-RU	GGACTTAAUTGTTGCCTGAGAAGATGGGC	asqJ deletion in A. nidulans
asqJdw-FU	GGCATTAAUTTCTTTAGAGATCTTCCTCCTAATACAG	asqJ deletion in A. nidulans
asqJdw-RU	GGTCTTAAUCGATAGATATTGTCGTTGATGGG	asqJ deletion in A. nidulans
AFpyrG-F	GGCATCGTCGAGGCTCTG	Construction of probe for southern blot
AFpyrG-R	GCTCGGTCGTTCGGCTG	Construction of probe for southern blot

Primers marked in blue are repeated.

S2 Text. NMR structural elucidations of niduclavin and niduporthin

The structures of niduclavin and niduporthin were established based on interpretations of 1D and 2D NMR data (^{1}H-NMR, ^{13}C-NMR, DQF-COSY, edHSQC, HMBC, H2BC, and NOESY).

Niduclavin

The DQF-COSY revealed five aromatic protons (H-5' to H-7'), expected to originate from phenylalanine (C-1' to C7' and NH), and a methyl substituted cyclic octaketide (C1 to C-16, 6-CH_3, 8-CH_3, and 14-CH_3). HMBC and DQF-COSY revealed a double bond between the α and β positions (C-2' and C-3') in the amino acid part of the molecule. The polyketide part of the molecule, was found to contain two diastereotopic CH_2-groups (C-7 and C-9), once again indicating a cyclic molecule. Four CH_3-groups (6-CH_3, 8-CH_3, 14-CH_3, and 16) and five CH-groups (C-4 to C-6, C8, C-10 to C-13, and C-15) were also identified. Analysis of the DQF-COSY and HMBC data established a decalin ring system including the linking of all four CH_3-groups to this part of the molecule. Furthermore, a W-coupling between H-15 and 14-CH_3 indicated the double bond between C-14 and C-15 to be in an *E*-configuration.

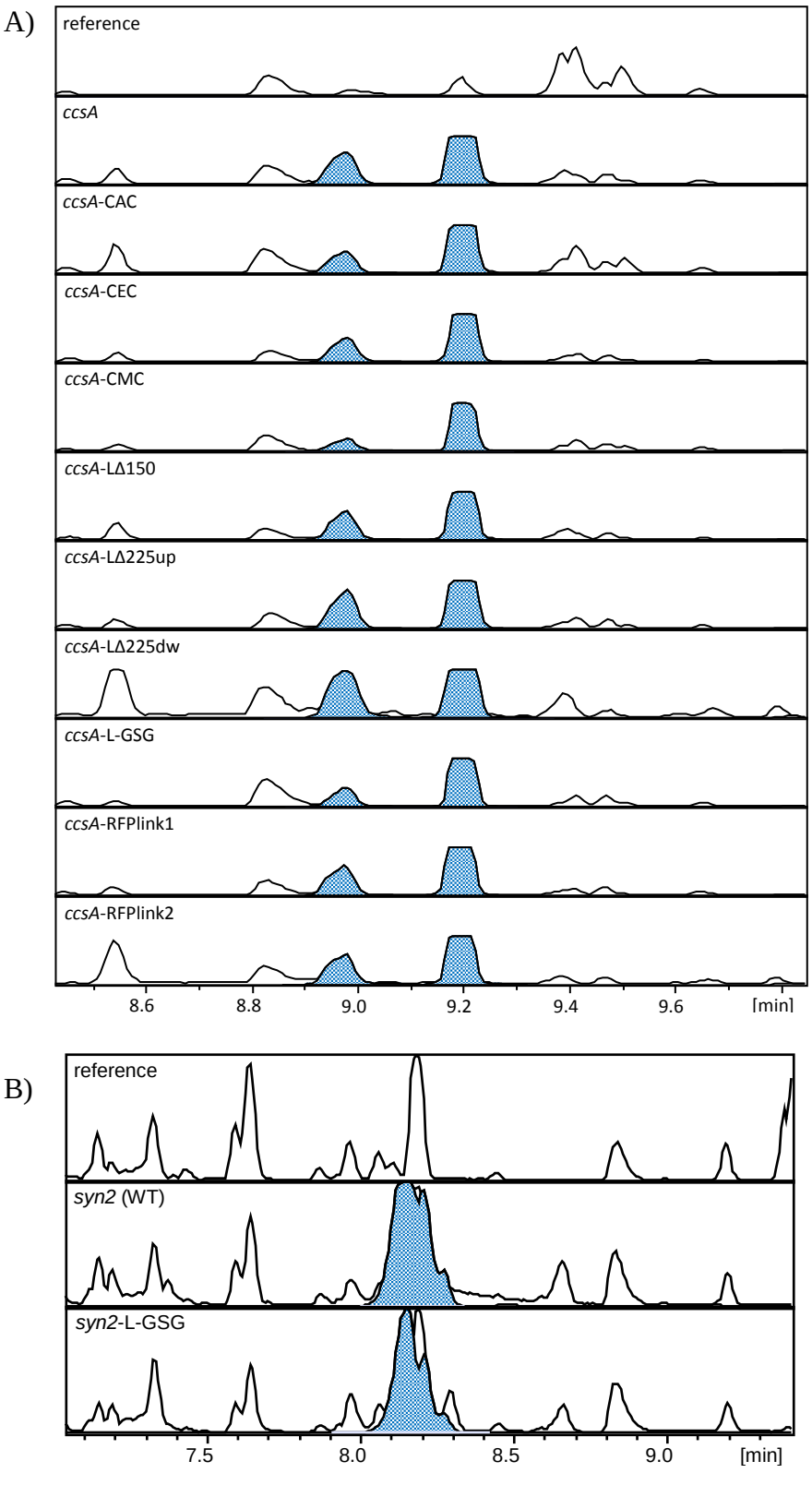
Assignment of relative stereochemistry of niduclavin was hindered by two protons with the same chemical shift (H-5 and H-9a, both at δ 1.75 ppm). As a result, it was not possible to assign the relative stereo chemistry to the two methyl groups (6-CH_3 and 8-CH_3).

Niduporthin

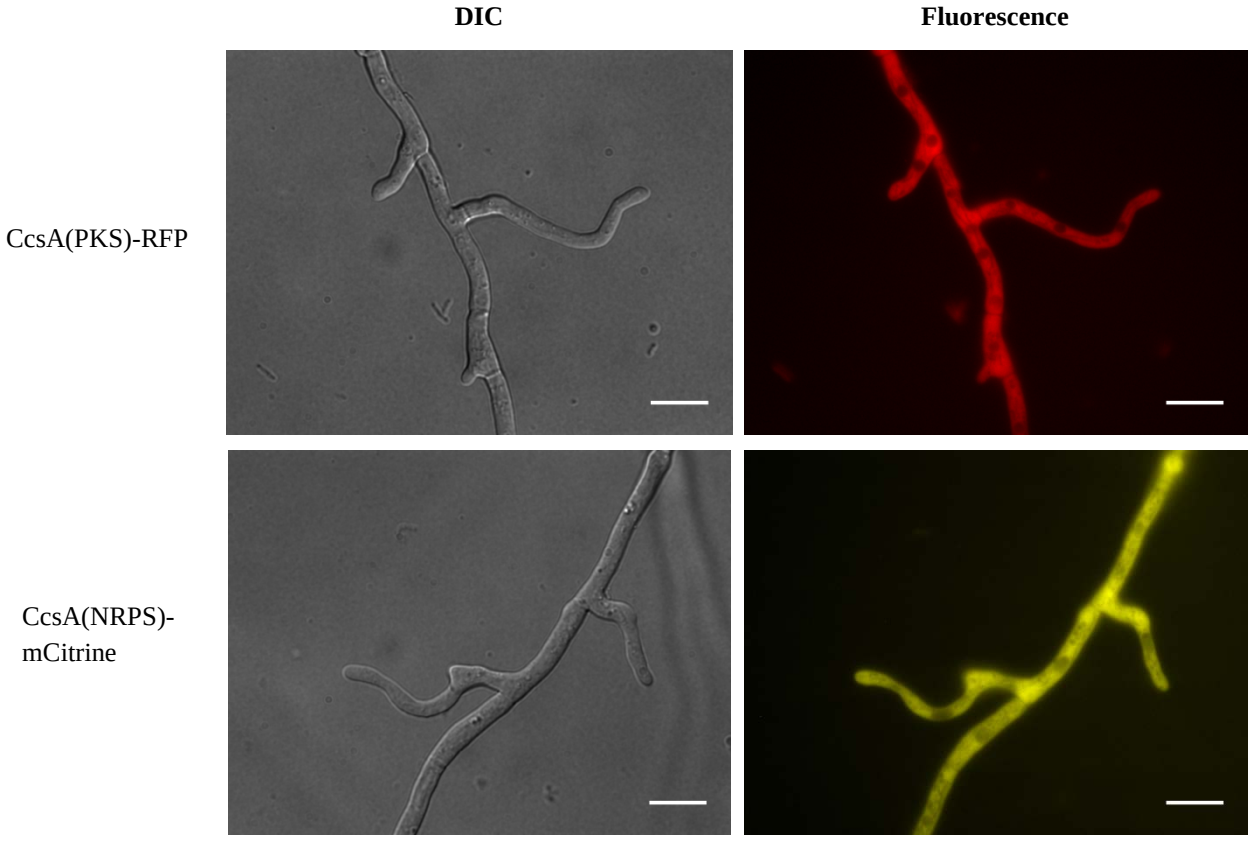
Downfield proton signals at 7-8 ppm were confirmed to belong to an indole heterocycle consisting of atoms C1'-C11' as well as NH-1 and NH-2, confirming the expected incorporation of a tryptophan moiety. However, HMBC and DQF-COSY also revealed that the link between the α and β positions (C-2' and C-3') had been converted into a double bond, a modification speculated to be the result of endogenous *A. nidulans* enzymes. Secondly, DQF-COSY and HMBC could be used to identify an octaketide (C-1 to C-16), containing two CH_3-groups (C-6 and C-16), three diastereotopic CH_2-groups (C-7, C-8 and C-9) and five CH-groups (C-4 to C-6 and C-10 to C-15). Considering the size of the spin system, as well as the placement of the two olefinic protons, H-11 and H-12, a decalin ring system formed via a [4+2] cycloaddition was proposed, also confirmed by the H2BC data. Two carbonyls were identified; one belonging to a ketone (C-3), and one to a tetramic acid derived lactam (C-1), both originating from the polyketide part of the molecule. HMBC showed correlations to the ketone (C-3) from both the decalin ring (H-4) and from the tetramic acid (H-1'), thereby linking the two parts of the molecule together.

The relative stereochemistry could be assigned based on the NOESY experiments, as well as biosynthetic/mechanistic considerations. NOESY correlations were observed between protons H-6 and H-13, and H-4 and H-10. This fits with the hypothesis that the decalin ring is expected to be formed via a [4+2] cycloaddition, leading to location of protons H-10 and H-13 on the same face, and the protons H-5 and H-10 placed on opposite face of the decalin ring respectively, due to the cycloaddition being intramolecular.

S3 Fig. Base peak chromatograms of *A. nidulans* extracts expressing various linker modified variants. Base peak chromatograms of A) *ccsA* and B) *syn2*. Highlighted areas represent EICs for A) niduclavin (*m/z* 416.2584), and B) niduporthin (*m/z* 427.2380).

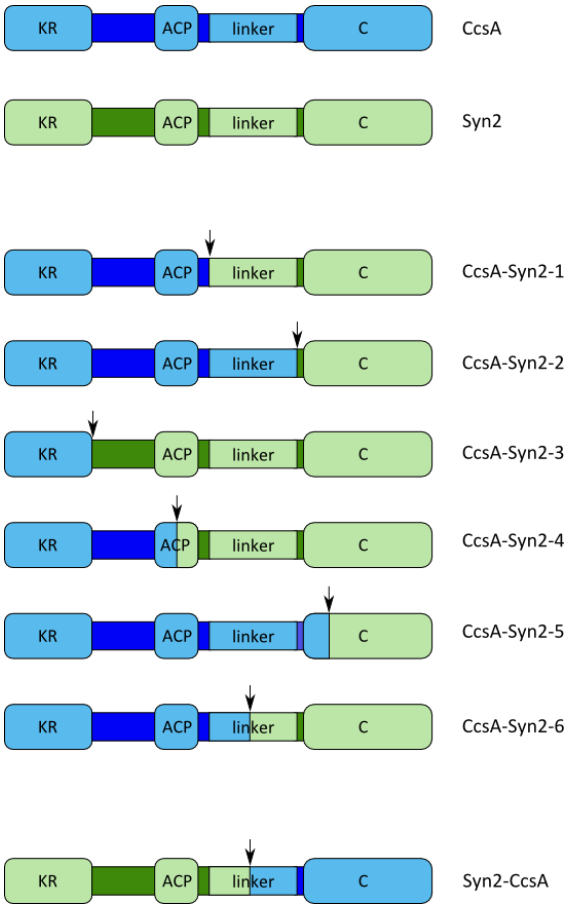


S4 Fig. Fluorescence tagging of individually expressed CcsA PKS- and NRPS modules. The RFP-tagged PKS module and the mCitrine-tagged NRPS module both appear to be localized to the cytoplasm. C-terminal RFP-tagging of the PKS module and C-terminal mCitrine-tagging of the NRPS module indicate cytoplasmic localization for both modules. N-terminal tagging of the modules revealed the same localization as seen for the C-terminal tagging (results not shown). Scale bar 10 µm. See also Supplemental Experimental Procedures.

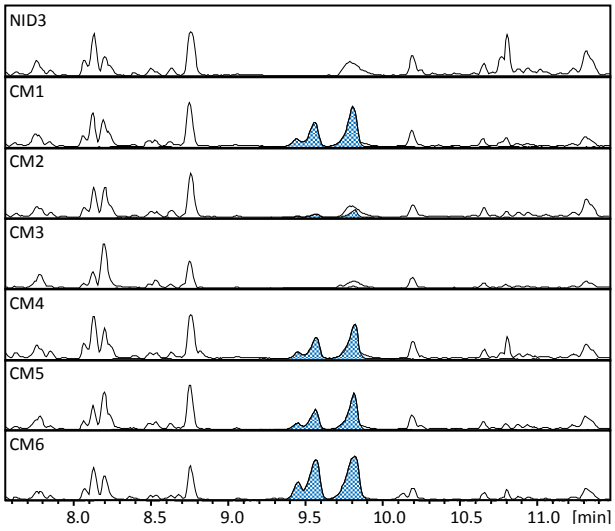


S5 Fig. Constructed fusions between CcsA and Syn2 PKS- and NRPS modules. A) Schematic illustration of the fusions between CccA and Syn2 PKS- and NRPS modules. Arrows indicate the point of fusion. B) Base peak chromatograms of *A. nidulans* extracts expressing CcsA-Syn2 chimeric PKS-NRPSs. Niduchimaeralin A elutes as several isomeric structures and are highlighted in blue (EIC @ *m/z* 455.2693).

A)



B)



Appendix 8

Transforming the lovastatin producing PKS, LovB, into a PKS-NRPS hybrid

Maria Lund Nielsen[1,2], Thomas Isbrandt[1,2], Uffe Hasbro Mortensen[1], Jakob Blæsbjerg Hoof[1], Mikael Rørdam Andersen[1,*] , and Thomas Ostenfeld Larsen[1,3,*]

Summary

The lovastatin nonaketide synthase LovB from *Aspergillus terreus* is a key enzyme in the biosynthesis of the blockbuster drug lovastatin. It is also one of the most well studied fungal polyketide synthases (PKSs). The presence of a C-terminal condensation (C) domain has raised the question as to whether LovB and fungal PKS-nonribosomal peptide synthetase (NRPS) hybrids share a common ancestor. However, the exact role of the C domain has remained uncertain since the domain does not serve the function of standard C domains, though some speculations have been about a possible Diels-Alderase-like activity. Here, we demonstrate engineering of the LovB PKS by fusing it to the NRPS module of the *A. clavatus* PKS-NRPS hybrid CcsA in order to obtain a novel PKS-NRPS hybrid product. The compatibility of the two modules suggests not only a common ancestor of LovB-type PKSs and PKS-NRPS hybrids, it also shows that the mechanisms for PKS-NRPS diversification are based on highly flexible synth(et)ases, which can produce compounds despite major modifications.

Keywords: Lovastatin, PKS-NRPS, Biosynthesis, Polyketides, *Aspergillus*

[1] Department of Biotechnology and Biomedicine, Technical University of Denmark, Søltofts Plads, 2800 Kongens Lyngby, Denmark

[2] Co-first author

[*] Corresponding authors: mr@bio.dtu.dk and tol@bio.dtu.dk

[3] Lead Contact

Introduction

Bioengineering of synthetic natural products is an attractive approach for producing novel bioactive compounds, which can be achieved by combining modules of enzymes involved in secondary metabolism. Polyketide synthases (PKSs) are major contributors to natural product diversity and this group of very complex enzymes has been the subject of extensive research to uncover the mechanisms for product formation (reviewed by several groups[1–3]). Best understood are the bacterial type I PKSs that work in an assembly-line fashion where the degree of polyketide reduction can often be predicted from the domain architecture of each module. However, the mechanisms of product formation in iterative type I highly reducing PKSs (hrPKSs), most typically found in fungi, are much less understood. For this type of PKS, polyketide synthesis proceeds through the repeated use of several catalytic domains, while exerting an intricate control of polyketide elongation, reduction, and methylation.[1]

The lovastatin nonaketide synthase LovB is one of the most extensively studied iterative hrPKSs. The product of LovB, dihydromonacolin L, is a precursor of lovastatin; a widely used cholesterol-lowering drug produced by *Aspergillus terreus*.[4,5] LovB stands out from many other fungal PKSs in that it contains a C-terminal condensation (C) domain homologous to domains found in nonribosomal peptide synthetases (NRPSs) and natural hybrids of PKSs and NRPSs (PKS-NRPSs), with the difference being the lack of a catalytically important His residue possibly being the reason for the lack of true C domain activity.[6] In fact, the presence of the C-terminal C domain in LovB has led to the speculation that LovB and PKS-NRPS hybrids share a common ancestor.[6,7] In support of this hypothesis is also the fact that the activity of LovB, like for many PKS-NRPSs, is dependent on the presence of a separate trans-acting enoyl reductase, LovC, for proper synthesis of dihydromonacolin L (Figure 1).[7,8] Boettger *et al*. conducted a phylogenetic analysis of the LovB C-terminal C domain and concluded that the domain is evolutionary divergent from the C domain of PKS-NRPSs and proposed a 'noncanonical' function for the LovB C domain.[6] Smith and co-workers[9] has suggested that the LovB C domain has adopted an alternative function, by catalyzing formation of the decalin ring through a [4+2] cycloaddition.[10,11] Removal of the C-terminal C domain has been shown to alter the programming of the enzyme, and instead of dihydromonacolin L, several truncated pyrones are produced.[7,9] This finding led to the suggestion that the C domain also catalyzed release of dihydromonacolin L from the enzyme, but this was later attributed to LovG, a thioesterase encoded in the lovastatin gene cluster.[12]

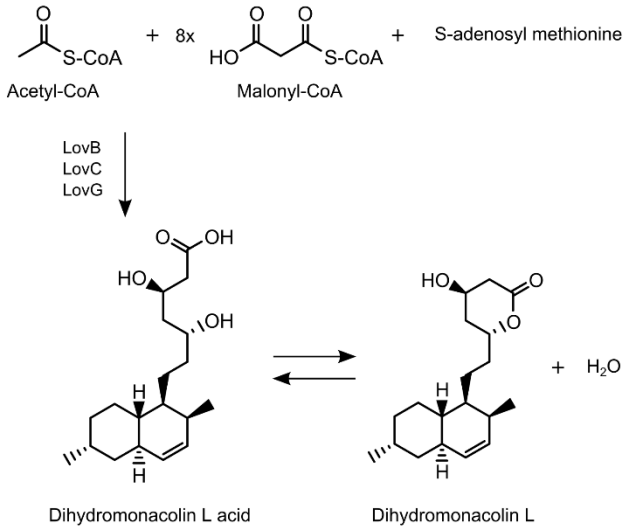


Figure 1. Biosynthesis of the lovastatin intermediate dihydromonacolin L. Synthesis proceeds through condensation of one acetyl-CoA starter unit and eight extender units of malonyl-CoA and is catalyzed by the nonaketide synthase LovB, the enoyl reductase LovC, and the thioesterase release factor LovG. Dihydromonacolin L acid is released from LovB and can undergo lactonization to yield dihydromonacolin L.

Boettger *et al*. previously investigated whether the supposedly lost PKS-NRPS activity of LovB could be restored.[6] Fusion of the LovB PKS with the NRPS module of the chaetoglobosin PKS-NRPS-encoding gene identified from *Penicillium expansum* (CAO91861)[13] did however not lead to synthesis of an amidated polyketide product. Following a number of bioinformatic analyses comparing LovB-type PKSs with hybrid PKS-NRPSs, they claimed an inherent incompatibility of the LovB C domain with the NRPS modules of hybrid PKS-NRPSs, and additionally that the C domain of LovB is required for the proper synthesis of dihydromonacolin L. Also Smith and co-workers attempted the construction of functional chimeric LovB variants by fusion of the LovB PKS module (without the C domain) to the NRPS module of the equisetin PKS-NRPS EqxS from *Fusarium heterosporum*. They were however also unsuccessful and did not yield any novel hybrid products.

In a previous study, we demonstrated that the modules of two PKS-NRPSs, sharing only 52 % sequence identity, could be successfully combined to produce the expected hybrid metabolites.[14] Our results led us to believe that structural similarity of the polyketide intermediate is a critical factor for construction of chimeric PKS-NRPS products, and that combining even more distantly related hybrids should be possible if their polyketide intermediates are structurally similar. Encouraged by these results we set out to attempt to revive the PKS-NRPS activity of LovB by fusing the PKS with the NRPS module from the cytochalasin E PKS-NRPS hybrid CcsA from *A. clavatus*.[15] Here, contrary to previous reports, we show that using an alternate strategy, LovB can indeed be combined with NRPS modules originating from PKS-NRPS hybrids to form a functional enzyme and thereby alter the activity of LovB to instead function as part of a PKS-NRPS hybrid. These findings support our earlier hypothesis that even distantly related PKS-NRPSs can be combined to form functional chimeric

enzymes, providing that a suitable substrate is produced by the PKS, suggesting that the mechanisms for PKS-NRPS diversification are based on substrate flexible synth(et)ases, able to produce new compounds from hybrid synthetases despite major modifications.

Results

Production of dihydromonacolin L acid in *A. nidulans*

As a reference, the activity of *A. terreus* LovB (ATEG_09961) in *A. nidulans* was tested by cloning the full LovB PKS with its C domain into an overexpression plasmid and transformed into the background strain NID1. Analysis by ultra-high performance liquid chromatography (UHPLC) coupled with diode array detection (DAD) and high-resolution mass spectrometry (HRMS) confirmed that expression of LovB alone is insufficient for synthesis of the lovastatin intermediate dihydromonacolin L, similar to what has previously been reported.[5] Furthermore, when co-expressing LovB with the trans-acting enoyl reductase LovC (ATEG_09963), UHPLC-DAD-HRMS revealed production of a compound with mass-to-charge ratio matching dihydromonacolin L acid (supporting information, figure S1). This result has also previously been reported in multiple studies.[7,10,16] The release factor LovG was not necessary for release of the PKS-NRPS product, indicating that *A. nidulans* holds a yet unidentified endogenous thioesterase activity capable of releasing the polyketide intermediate.

Design of LovB-CcsA chimeras

We next constructed two fusions of LovB and *A. clavatus* CcsA (ACLA_078660) and expressed them heterologously in *A. nidulans* (Figure 2A). For mutant 1 (Figure 2B), the compatibility of the LovB C domain with the remaining domains of the CcsA NRPS module was tested by extending the full-length LovB with the remaining downstream NRPS domains, *i.e.* the adenylation (A) domain, the peptide carrier protein (PCP) domain, and the reductase (R) domain. For mutant 2 (Figure 2C), the LovB C domain was replaced with the full-length NRPS module of CcsA. If the LovB C domain has indeed diverged and lost its condensation activity, we hypothesized that this could be bypassed by instead replacing the C domain with a full-length NRPS module, thus creating a PKS-NRPS fusion.

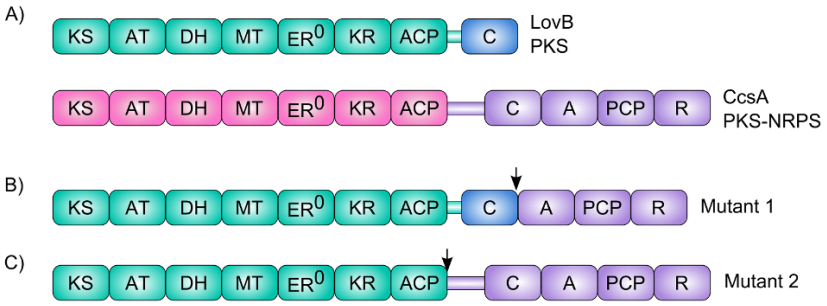


Figure 2. Construction of chimeric PKS-NRPS hybrids by fusion of LovB and CcsA modules. A) The lovastatin nonaketide synthase LovB from A. terreus comprise the typical domains found in iterative hrPKSs but includes also a C-terminal C domain. CcsA encodes a hybrid PKS-NRPS that is involved in the biosynthesis of cytochalasin E in A. clavatus. Two LovB-CcsA chimeric hybrids were constructed. B) In mutant 1 LovB was extended with the remaining domains of the CcsA NRPS module – the A, PCP, and R domains. C) In mutant 2 the LovB C domain was replaced with the full-length NRPS module of CcsA. The long linker of the CcsA PKS-NRPS hybrid was preserved in this construct to prevent possible restraints of the shorter LovB linker. ER^0 denotes an inactive enoyl reductase domain. Arrows indicate transition points between the two PKS-NRPSs. See also Figure S1 and Table S3.

In our efforts to induce a PKS-NRPS activity of LovB we selected the NRPS module of the *A. clavatus* cytochalasin PKS-NRPS for a number of reasons: In a previous study, we have shown that the *A. clavatus* PKS-NRPS CcsA can be functionally expressed in *A. nidulans*, resulting in production of a highly reduced octaketide product.[14] In the same study, we also successfully swapped the PKS- and NRPS modules of CcsA and investigated how modules can be combined without disrupting enzyme activity. Furthermore, since it has been shown that the removal of the native LovB C domain results in production of truncated pyrones that do not contain the decalin ring,[7,9] we assumed that the most compatible NRPS module would be one that normally accepts shorter polyketide intermediates, *e.g.* heptaketides or octaketides.

The LovB C domain demonstrates incompatibility with the A-PCP-R domains of CcsA

Mutant 1 was constructed by fusion of the *A. terreus lovB* with the *A. clavatus ccsA* fragment encoding the A-PCP-R domains (Figure 2B). The exact transition point of the two sequences was decided based on a sequence alignment of *lovB* and *ccsA* using Clustal Omega (Figure S2). Since the expression of LovC is necessary for the activity of LovB, we transformed the *lovB-ccsA* fusion into an *A. nidulans lovC* overexpression strain. The resulting strain was analyzed by UHPLC-DAD-HRMS, and the metabolite profile of mutant 1 revealed the production of dihydromonacolin L acid, suggesting that LovB activity was not affected by the presence of the three NRPS domains of CcsA. We were furthermore unable to identify any new nitrogen-containing polyketide products in the extracts of mutant 1. In fact, the metabolite profile of mutant 1 was identical to the *lovB*/*lovC* expressing strain. Hence, these results are comparable to the results of Hertweck and co-workers who were also unable to restore the condensation activity of the native C domain of LovB when they fused it to the NRPS module of the chaetoglobosin PKS-NRPS from *P. expansum*.[6] The native C domain thus seems to be robust in conferring incompatibility to downstream NRPS domains.

The PKS-NRPS activity of LovB can be restored by fusion to the CcsA NRPS module

The expression construct for mutant 2 was assembled by fusing the *lovB* PKS-encoding fragment with the NRPS-encoding fragment of *ccsA*. The presence of a long highly variable intermodular region of approximately 150 amino acids between the PKS- and NRPS modules of CcsA was identified after prediction of the ACP- and C domains of CcsA using CDD from NCBI,[17] based on our previous work.[14] An equivalent but shorter region (approximately 60 amino acids) was predicted in LovB. To ensure that sufficient distance was put between the two modules, and to prevent possible restraints by the intermodular linker, the long CcsA linker was preserved for construction of mutant 2 (see Figure 2C and Figure S2).

Interestingly, the metabolite profile of mutant 2 revealed the presence of two major isomeric compounds with a mass of 387.2205 Da ($[M+H]^+$ = 388.2278), corresponding to the elemental compositions $C_{26}H_{30}NO_2$ (Figure 3). The predicted elemental composition revealed the presence of a nitrogen atom, indicating that mutant 2 did in fact express a functional chimeric PKS-NRPS presumably by incorporating a phenylalanine residue into a LovB derived polyketide intermediate. Assuming incorporation of phenylalanine, our results also indicated that the polyketide elongation by LovB had arrested one cycle short of normal nonaketide biosynthesis, to instead yield an octaketide.

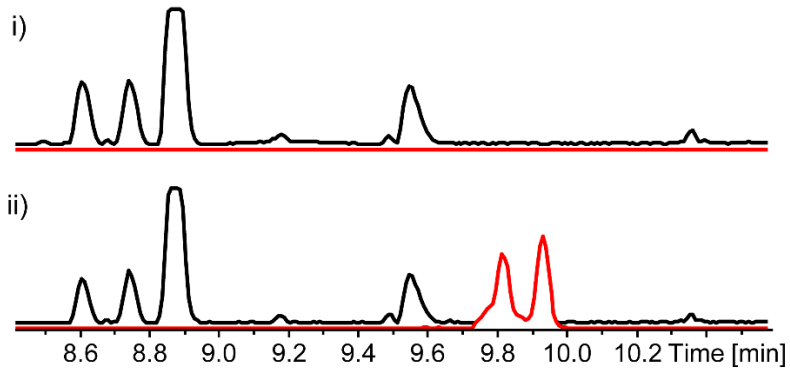


Figure 3. Investigating the compatibility of LovB with the NRPS module of CcsA in A. nidulans. Trace i: Overlaid BPC (black line) and EIC (red line) of the A. nidulans background strain .Trace ii: Overlaid BPC (black line) and EIC (red line) of mutant 2 showing production of isomers of the novel compound ($[M+H]^+$ = 388.2271±0.001 Da), terreclavin.

The structure of the major isomer of the novel hybrid compound (*m/z* = 388.2278, $C_{26}H_{30}NO_2$) was determined by use of one and two-dimensional NMR spectroscopy (Table S1 and Table S2). The data revealed a compound containing five aromatic protons, nine alkenes, and two CH_3 groups. From the multiplicity edited HSQC, three diastereotropic CH_2 groups could be identified, along with the two CH_3 groups and 15 CH groups of which five were aromatic. HMBC correlations linked the aromatic ring to a tetramic acid moiety, similar to what has previously been observed in other PKS-NRPS hybrid products.[14] Using DQF-COSY, H2BC, and HMBC the

polyketide was confirmed to be a linear highly reduced octaketide with a single methylation at C-7, as also seen in natural LovB products. HMBC couplings to C-14 from H-1', H-11, and H-12 linked the tetramic acid and the polyketide. NOESY correlations between H-1' and H-3' determined the configuration of the double bond between C-1' and C-2'. Similarly, NOESY correlations between H-2 and H-4, H-3 and H-5, H-10 and H-12, and between H-11 and H-13 indicated the polyketide alkenes for the major isolated isomer to all be in *trans* configurations, whereas we speculate the other isomer to contain one or several *cis* double bonds. The structure of the compound, which we have name terreclavin, is shown in Figure 4.

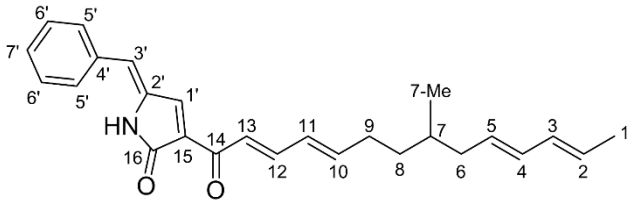


Figure 4. Structure of the major novel hybrid product terreclavin produced by mutant 2 (exact mass 387.2198 Da). The polyketide is an octaketide, and thus one extension shorter than the native LovB products found in A. terreus. See also Table S1.

In summary, the structure suggests that terreclavin is biosynthesized by the LovB-CcsA chimeric enzyme (mutant 2) through seven elongation cycles using malonyl-CoA extender units to achieve production of the octaketide. The polyketide is subsequently transferred to the NRPS module that attaches the phenylalanine residue and is finally released in a reductive manner yielding the tetramic acid moiety (Figure 5).

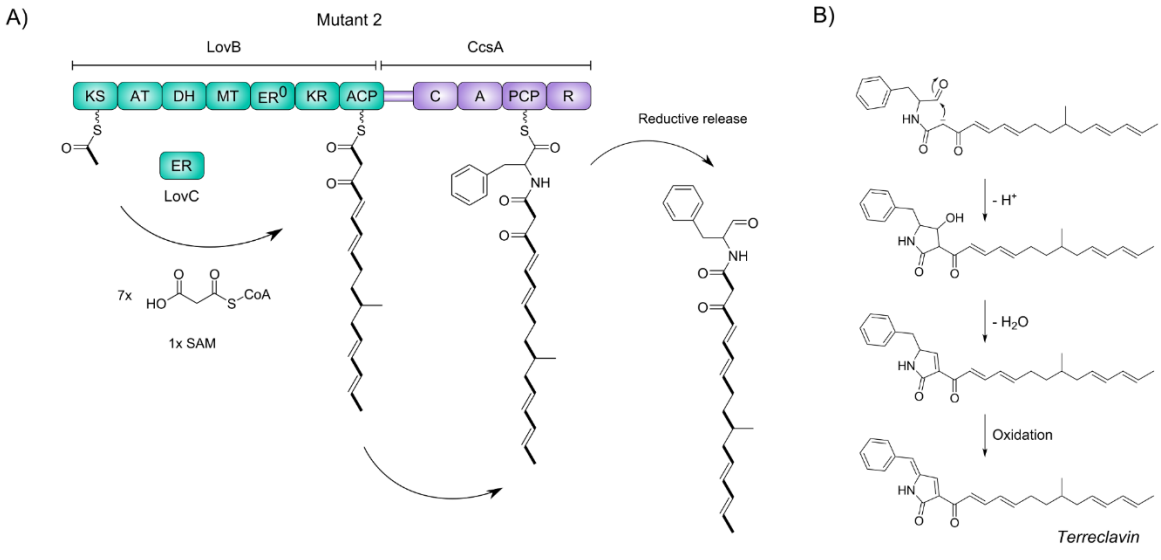


Figure 5. Proposed biosynthesis of terreclavin by the LovB-CcsA chimeric enzyme (mutant 2). A) Synthesis of the terreclavin polyketide backbone proceeds through condensation of an acetyl starter unit and seven malonyl-CoA extender units. These reactions are catalyzed by the LovB type I hrPKS and the enoyl reductase LovC. Attachment of the phenylalanine residue and subsequent release of the polyketide-nonribosomal peptide product are mediated by the CcsA NRPS module, which is C-terminally fused to LovB. B) Proposed mechanism for the cyclisation (as also suggested by Fujii et al., 2013) and oxidation/desaturation of terreclavin. The desaturation step is likely the result of an endogenous A. nidulans activity.

Discussion

In the extracts of mutant 1 we were unable to detect any novel hybrid products, indicating that the LovB C domain as expected was unable to form functional interactions with the downstream NRPS domain of CcsA. Boettger *et al.* conducted a similar experiment where LovB was fused to the A-PCP-R domains of the chaetoglobosin PKS-NRPS CheA from *P. expansum*, and they were equally unsuccessful in establishing a functional interaction between the two non-cognate modules in this LovB-CheA fusion.[6] From the structure of the *Bacillus subtilis* terminal NRPS module SrfA-C, it has been shown that the C domain and the major part of the A domain together constitute a catalytic platform onto which the PCP and an A subdomain rearrange to facilitate catalysis.[19] Recently, two additional crystal structures of holo-NRPS modules were described.[20] These structures confirmed that the C- and A- domains share a large interface, and work in a concerted fashion simultaneously adopting their catalytic conformations. The critical partnership of the C- and A domains was also recently substantiated when it was shown that the monospecificity of an NRPS module required the presence of cognate C- and A domains, and hence that the C domain also exhibits a specificity-regulatory role.[21] We speculate that variation of the LovB C domain, i.e. the lack of the catalytically crucial His residue, prevents a functional association between the C domain of LovB and the A domain of CcsA in mutant 1.

The exact catalytic activity and mechanism of the native C domain of LovB is still not fully understood. It has previously been shown that the presence of the LovB C domain is critical for correct assembly of the polyketide backbone and that this domain is involved in controlling proper chain length.[7,9] It is however unlikely that the "conventional" C domains found in PKS-NRPSs serve this function. In two previous cases the PKS- and NRPS modules of hybrids have successfully been swapped despite the fact that the polyketide products were of different sizes. In one study, the PKS module of the equisetin PKS-NRPS (EqxS) was combined with the NRPS module of the fusaridione A PKS-NRPS (FsdS), both from *F. heterosporum*.[9] In another study, the dissected PKS modules of the cyclopiazonic acid PKS-NRPS (CpaS) of *A. flavus* was expressed in *Saccharomyces cerevisiae* with the equally dissected NRPS module of the aspyridone NRPS module (ApdA) from *A. nidulans*. In both studies, the successful swaps yielded the expected hybrid product, which showed that the programming of the PKS module in regard to the polyketide chain length is not affected by fusion to a non-cognate module. Thus, the C domains of PKS-NRPS hybrids does not appear to be involved in chain length determination; a conclusion, which is also in agreement with a study by Lazarus and co-workers.[22] All in all, it appears that the LovB C domain has adopted an alternative function different from the function of conventional NRPS C domains, but a function that is nonetheless essential for proper synthesis of dihydromonacolin L.

Mutant 2 represents the first successful example of a functional chimeric PKS-NRPS that has been constructed by adding an entire NRPS module to a PKS. In a previous study we demonstrated the flexibility in the length and sequence of the intermodular linker of PKS-NRPSs.[14] Our results, as well as the successful module swapping in previous reports.[9,23,24] have also shown a flexibility in NRPS modules, *i.e.* that NRPS modules are able to catalyze condensation of amino acids to more than one polyketide substrate. Boettger *et al.* speculated that the evolutionary divergence of LovB from the CheA PKS-NRPS was to account for the lack of PKS-NRPS activity when fusing the LovB PKS moiety with the NRPS module of CheA.[6,25] However, our results clearly demonstrate that the lack of function was not due to a general incompatibility of LovB with NRPS modules from PKS-NRPSs, which underlines the evolutionary relationship between PKS-NRPS hybrids and the LovB-type PKSs.

In light of the two previous unsuccessful attempts of inducing PKS-NRPS activity of LovB,[6,9] our results also highlight the challenges still faced by the scientific community in predicting what is required to obtain functional chimeric PKS-NRPSs. The putative CcsA, CheA, and EqxS polyketide intermediates are structurally similar. CcsA, CheA, and EqxS are all distantly related to LovB with the CcsA PKS-C fragment displaying only slightly higher identity to LovB (39 %) as compared to CheA-C versus LovB (33 %), and EqxS-C versus LovB (36%). Hence, the exact reasons for their lack of success are not immediately obvious. However, it is not possible to rule out problems such as functional expression of heterologous genes, *i.e*. expression of the *P. expansum cheA* fragment in *A. terreus*, or the expression of *lovB* in *F. heterosporum*.

Our finding that LovB and CcsA, two distantly related enzymes, can be combined to form functional chimeric enzymes does indeed seem to agree with our previous hypothesis that high sequence identity is not a critical factor for the design of functional PKS-NRPS hybrids. Instead, we hypothesize that the structures of the polyketide intermediate is the primary determining factor for production of “chimeric” metabolites. Also, in previous studies where successful PKS-NRPS module swaps have been accomplished, the length of the PKS intermediates of the two enzymes involved differ only by up to a single extension.[9,14,22] Thus, despite the above-mentioned unsuccessful swaps, it appears that for construction of functional PKS-NRPS hybrids, it is appropriate to consider the polyketide structures of the compounds to be combined. However, it must be stresses that further combinatorial studies of PKS-NRPS hybrids are needed to corroborate this hypothesis.

The presence of phenylalanine in terreclavin isolated from mutant 2 show the successful engineering of PKS-NRPS activity of the expressed LovB-CcsA hybrid. The chain length of the polyketide moiety is shorter compared to what is seen for natural LovB products. Further, the PKS moiety also has a higher level of reduction compared to what in seen during lovastatin biosynthesis (Figure 6). A double bond is found between

C-12 and C-13, where normally a single bond would be, and C-14 and C-16 are both completely non-reduced, which could suggest a possible interaction of the LovB C domain under normal conditions. The position of the methylation, however, remains at carbon 7, which is also the case for natural LovB derived compounds like lovastatin and dihydromonacolin analogs.

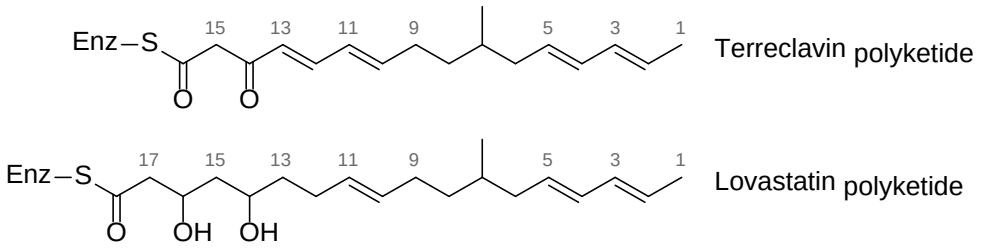


Figure 6. Comparison of the terreclavin and lovastatin polyketide backbone, illustrating difference in length and reduction.

Based on the structure of terreclavin, an unexpected octaketide has been fused with a phenylalanine moiety. We speculate that an additional function of the LovB C domain, could be influencing the chain elongation of the PK. By forcing an additional extension by keeping the product in the PKS machinery, the resulting product in the natural LovB PKS is a nonaketide. This raises the question whether the true product of the LovB PKS is in fact an octaketide, and that the C domain is somehow able to increase the chain length of the final product. Further experiments are needed to provide a more clear answer to this matter.

It is noteworthy that terreclavin contains a linear polyketide chain and not the usual decalin ring structure seen in lovastatin, despite that the double bond pattern of the polyketide chain would allow for a [4+2] cycloaddition. We speculate that this could be due to conformational effects during biosynthesis, caused by the attached foreign NRPS module, or as a result of the double bond between C-12 and C-13. Furthermore, we note that a double bond is present between carbons C-2' and C-3' in terreclavin, a feature previously seen when expressing *ccsA* in *A. nidulans* [14]. This further strengthens our earlier hypothesis that this modification is a result of an endogenous *A. nidulans* activity.

Phylogenetic studies suggest a common origin of PKS-NRPSs.[26] In fact, in an earlier study it was shown that the NRPS A domain of fungal PKS-NRPSs is phylogenetically distinct when compared to other fungal PKSs and NRPSs.[27] If LovB has evolved from PKS-NRPS hybrids, this could explain the compatibility of LovB and CcsA. Yet the question then remains whether it is possible to combine NRPS modules from PKS-NRPSs to PKSs not evolutionary related to the LovB-type PKS.

As we have demonstrated here, it is possible to increase chemodiversity by module swapping of PKS-NRPSs by genetic engineering, and thus, it seems also plausible that chemical diversification arise in nature by shuffling

of PKS- and NRPS-encoding sequences both through genome shuffling and horizontal gene transfer. A possible model for the evolution of PKS-NRPS hybrids and the LovB-type PKSs, which is consistent with all available data, is presented in Figure 7. We propose that diversification can occur through several different events: Fusion of independent PKSs with NRPSs (Figure 7A). Whether this has happened more than the one event, is unknown. Other events include exchange of PKS- or NRPS modules in existing PKS-NRPS hybrids (Figure 7B), or separation of PKS-NRPS hybrids leading to PKSs and NRPSs that have become functional as individual enzymes (Figure 7C). Our present results support earlier hypotheses that LovB-type PKSs evolved from an early PKS-NRPS by truncation and loss of the C-terminal A-PCP-R domains, followed by C domain differentiation (Figure 7D). By fusion of the LovB PKS module with an NRPS module from a PKS-NRPS hybrid, thereby creating a novel chimeric enzyme, we have shown that it is possible to restore the PKS-NRPS activity of LovB (Figure 7E).

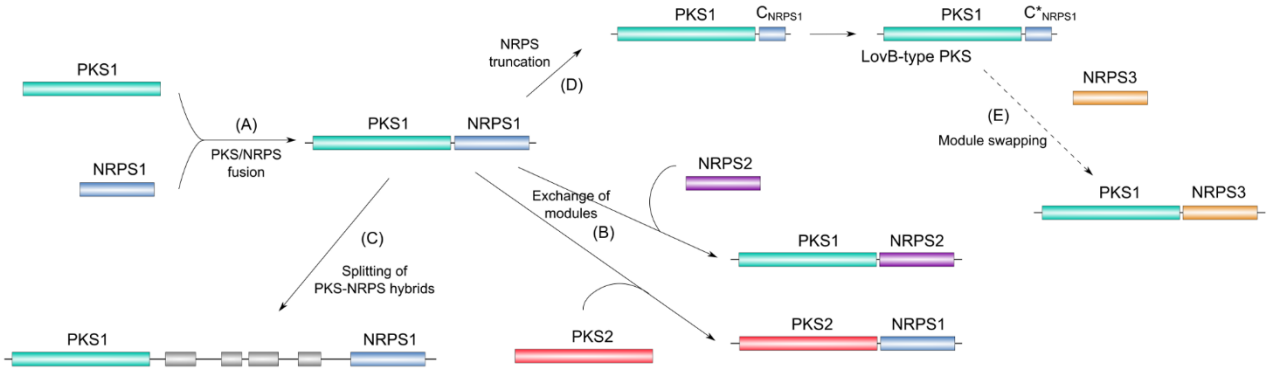


Figure 7. Model for the development and diversification of PKS-NRPS hybrids as well as the LovB type PKSs. PKS-NRPSs are believed to have arisen through fusion of single PKS- and NRPS modules (A). Subsequent diversification emerges either from PKS- and NRPS module exchange (B) or through splitting of existing PKS-NRPS hybrids (C). In the case of LovB-type PKSs, we believe the A-PCP-R domains were lost and that the C domain subsequently differentiated and adopted a new function (D). By fusion of the LovB PKS module with an NRPS from a PKS-NRPS hybrid the PKS-NRPS activity of LovB was restored (E).

Significance

This study demonstrates the successful fusion of two PKS/NRPS proteins with the generation of a previously unknown compound. Interestingly, the apparent flexibility of natural PKS-NRPS hybrids could provide a mechanistic insights to the remarkable chemical diversity observed for natural products of mixed biosynthetic origin. The construction of a functional LovB-NRPS fusion also indicates the evolutionary origin of the LovB synthase as a “malfunctioning” truncation. Thus, for all of the events described in Figure 7, we work from the hypothesis that if a given functional recombination is possible with our crude methods in the laboratory, it will also be possible through natural recombination events. This allows us both to build the model for PKS-NRPS

diversification as demonstrated above, but also it is an optimistic prospect for future synthetic biology effort in generating new biologically active hybrid compounds.

Experimental procedures

Strains and Media

Heterologous expression was performed in *A. nidulans* strain NID1 (*argB2, pyrG89, veA1, nkuAΔ*) strain IBT 29539.[28] Genes were amplified from genomic DNA of *A. terreus* NIH 2624 and *A. clavatus* NRRL1, which was extracted using the FastDNA™ SPIN Kit for Soil DNA extraction (MP Biomedicals, USA). Plasmid construction- and propagation was achieved using *Escherichia coli* strain DH5α.

A. nidulans was grown on solid and liquid minimal medium (MM), solid MM supplemented with 1.3 mg/ml 5-fluoroorotic acid (5-FOA), and solid transformation medium as described by Nødvig *et al*..[29] Media were supplemented with 4 mM arginine and when needed 10 mM uracil and 10 mM uridine. *E. coli* DH5α was grown in Luria-Bertani (LB) medium (10 g/l tryptone (Bacto), 5 g/l yeast extract (Bacto), 10 g/l NaCl (pH 7.0)) supplemented with 100 μg/ml ampicillin.

Vector construction

DNA fragments were generated using the PfuX7 polymerase[30] and primers from Integrated DNA Technology, Belgium (Table S3). Vectors were constructed by Uracil-Specific Excision Reagent (USER) fusion of PCR fragment into compatible expression plasmids. All plasmids contained a *Pac*I/*Nt.Bbv*CI USER cassette, the constitutive promoter P*gpdA*, the T*trpC* terminator, *A. fumigatus pyrG* for selection and 2 kb up- and downstream targeting sequences for integration into specific targeting sites in the *A. nidulans* genome[31]. *pyrG* is flanked by direct repeats for counter selection on 5-FOA. The plasmids also contain the ampicillin resistance gene for selection in *E. coli*. Plasmids were purified using the GenElute™ Plasmid Miniprep Kit (Sigma-Aldrich), and verified by restriction analysis. Prior to transformation, plasmids were linearized with *Swa*I, and protoplastation, transformation and strain validation of *A. nidulans* strains was performed as described by Nødvig *et al*..[29] A list of strains used in this study is provided in Table S4.

Chemical analysis

Strains of *A. nidulans* were grown for 6 days at 37°C on solid MM with all necessary supplements. The *A. nidulans* strain NID3 was used as a reference strain (Table S4). Plug extractions were performed as described by Smedsgaard[32] with the exception that metabolites were extracted with 3:1 ethylacetate:isopropanol

containing 1 % formic acid. The samples were analyzed on a maXis 3G orthogonal acceleration quadrupole time-of-flight mass spectrometer (Bruker Daltonics) equipped with an electrospray ionization (ESI) source and connected to an Ultimate 3000 UHPLC system (Dionex), equipped with a Kinetex 2.6 µm C18, 100mm x 2.1 mm column (Phenomenex). The method applied was described by Holm *et al*..[33]

Metabolite purification

For purification of terreclavin, mutant 1 strain was inoculated in 10x200 ml semi-liquid CYA medium supplemented with 4 mM arginine. After 7 days, the media was separated from the biomass, and both were extracted two times with ethyl acetate (EtOAc). The crude extract was subjected to flash column chromatography on an Isolera One system (Biotage, Uppsala, Sweden), using a diol column and stepwise elution with 100 % heptane; 1:1 heptane/dichloromethane (DCM); 100 % DCM; 1:1 DCM/EtOAc; 100 % EtOAc; 1:1 EtOac/MeOH; and 100 % MeOH. The fractions containing the compound of interest were further purified by semi-preparative reversed phase LC on a Waters 600 controller equipped with a 996 photodiode array detector using a Kinetex Core-Shell C18 column (Phenomenex, Torrance, California, USA).

NMR analysis

NMR spectra were recorded on a Bruker Avance 800 MHz spectrometer at the Department of Chemistry, at the Technical University of Denmark. Spectra were acquired using standard pulse sequences. The deuterated solvent was CD_3OD and signals were referenced by solvent signals for CD_3OD at $\delta H = 3.31$ ppm and $\delta C = 49.00$ ppm. The NMR data was processed in TopSpin 3.5pl5 (Table S2). Chemical shifts are reported in ppm (δ) and scalar couplings are reported in hertz (Hz) (Table S1). The sizes of the J coupling constants in the table are the experimentally measured values from the 1D ^1H and DQF-COSY spectra. There are minor variations in the measurements, which may be explained by the uncertainty of J and the spectral digital resolution.

Author contributions

Conceptualization, M.L.N., T.I., U.H.M., J.B.H., M.R.A., T.O.L.; Formal Analysis, M.L.N., T.I.; Investigation, M.L.N., T.I.; Writing – Original Draft, M.L.N., T.I.; Writing – Review & Editing, M.L.N., T.I., J.B.H., M.R.A., T.O.L.; Funding Acquisition, T.O.L.

References

(1) Weissman, K. J. *Nat. Prod. Rep.* 2015, *32* (3), 436.

(2) Hertweck, C. *Trends Biochem. Sci.* 2015, *40* (4), 189.

(3) Robbins, T.; Liu, Y.-C.; Cane, D. E.; Khosla, C. *Curr. Opin. Struct. Biol.* 2016, *41*, 10.

(4) Hendrickson, L.; Davis, C. R.; Roach, C.; Nguyen, D. K.; Aldrich, T.; McAda, P. C.; Reeves, C. D. *Chem. Biol.* 1999, *6* (7), 429.

(5) Kennedy, J. *Science (80-.).* 1999, *284* (5418), 1368.

(6) Boettger, D.; Bergmann, H.; Kuehn, B.; Shelest, E.; Hertweck, C. *Chembiochem* 2012, *13* (16), 2363.

(7) Ma, S. M.; Li, J. W.-H.; Choi, J. W.; Zhou, H.; Lee, K. K. M.; Moorthie, V. a; Xie, X.; Kealey, J. T.; Da Silva, N. a; Vederas, J. C.; Tang, Y. *Science (80-.).* 2009, *326* (5952), 589.

(8) Halo, L. M.; Marshall, J. W.; Yakasai, A. A.; Song, Z.; Butts, C. P.; Crump, M. P.; Heneghan, M.; Bailey, A. M.; Simpson, T. J.; Lazarus, C. M.; Cox, R. J. *ChemBioChem* 2008, *9* (4), 585.

(9) Kakule, T. B.; Lin, Z.; Schmidt, E. W. *J. Am. Chem. Soc.* 2014, *136* (51), 17882.

(10) Auclair, K.; Sutherland, A.; Kennedy, J.; Witter, D. J.; Van den Heever, J. P.; Hutchinson, C. R.; Vederas, J. C. *J. Am. Chem. Soc.* 2000, *122*, 11519.

(11) Witter, D. J.; Vederas, J. C. *J. Org. Chem* 1996, *61* (i), 2613.

(12) Xu, W.; Chooi, Y. H.; Choi, J. W.; Li, S.; Vederas, J. C.; Da Silva, N. a.; Tang, Y. *Angew. Chemie* 2013, *52* (25), 6472.

(13) Schümann, J.; Hertweck, C. *J. Am. Chem. Soc.* 2007, *129* (31), 9564.

(14) Nielsen, M. L.; Isbrandt, T.; Petersen, L. M.; Mortensen, U. H.; Andersen, M. R.; Hoof, J. B.; Larsen, T. O. *PLoS One* 2016, *11* (8), e0161199.

(15) Qiao, K.; Chooi, Y. H.; Tang, Y. *Metab. Eng.* 2011, *13* (6), 723.

(16) Sorensen, J. L.; Vederas, J. C. *Chem. Commun.* 2003, No. 5, 1492.

(17) Marchler-Bauer, A.; Derbyshire, M. K.; Gonzales, N. R.; Lu, S.; Chitsaz, F.; Geer, L. Y.; Geer, R. C.; He, J.; Gwadz, M.; Hurwitz, D. I.; Lanczycki, C. J.; Lu, F.; Marchler, G. H.; Song, J. S.; Thanki, N.; Wang, Z.; Yamashita, R. A.; Zhang, D.; Zheng, C.; Bryant, S. H. *Nucleic Acids Res.* 2015, *43* (D1), D222.

(18) Fujii, R.; Minami, A.; Gomi, K.; Oikawa, H. *Tetrahedron Lett.* 2013, *54* (23), 2999.

(19) Tanovic, A.; Samel, S. A.; Essen, L.-O.; Marahiel, M. A. *Science (80-.).* 2008, *321* (5889), 659.

(20) Drake, E. J.; Miller, B. R.; Shi, C.; Tarrasch, J. T.; Sundlov, J. A.; Leigh Allen, C.; Skiniotis, G.; Aldrich, C. C.; Gulick, A. M. *Nature* 2016, *529* (7585), 235.

(21) Meyer, S.; Kehr, J.; Mainz, A.; Dehm, D.; Petras, D.; Süssmuth, R. D.; Dittmann, E. *Cell Chem. Biol.* 2016,

23 (4), 462.

(22) Heneghan, M. N.; Yakasai, A. a.; Williams, K.; Kadir, K. a.; Wasil, Z.; Bakeer, W.; Fisch, K. M.; Bailey, A. M.; Simpson, T. J.; Cox, R. J.; Lazarus, C. M. *Chem. Sci.* 2011, *2* (5), 972.

(23) Xu, W.; Cai, X.; Jung, M. E.; Tang, Y. *J. Am. Chem. Soc.* 2010, *5* (C), 13604.

(24) Fisch, K. M.; Bakeer, W.; Yakasai, A. a; Song, Z.; Pedrick, J.; Wasil, Z.; Bailey, A. M.; Lazarus, C. M.; Simpson, T. J.; Cox, R. J. *J. Am. Chem. Soc.* 2011, *133* (41), 16635.

(25) Boettger, D.; Hertweck, C. *Chembiochem* 2013, *14* (1), 28.

(26) Bushley, K. E.; Turgeon, B. G. *BMC Evol. Biol.* 2010, *10* (1), 26.

(27) Lawrence, D. P.; Kroken, S.; Pryor, B. M.; Arnold, A. E. *PLoS One* 2011, *6* (11), e28231.

(28) Nielsen, J. B.; Nielsen, M. L.; Mortensen, U. H. *Fungal Genet. Biol.* 2008, *45* (3), 165.

(29) Nødvig, C. S.; Nielsen, J. B.; Kogle, M. E.; Mortensen, U. H. *PLoS One* 2015, *10* (7), e0133085.

(30) Nørholm, M. H. H. *BMC Biotechnol.* 2010, *10*, 21.

(31) Hansen, B. G.; Salomonsen, B.; Nielsen, M. T.; Nielsen, J. B.; Hansen, N. B.; Nielsen, K. F.; Regueira, T. B.; Nielsen, J.; Patil, K. R.; Mortensen, U. H. *Appl. Environ. Microbiol.* 2011, *77* (9), 3044.

(32) Smedsgaard, J. *J. Chromatogr. A* 1997, *760* (2), 264.

(33) Holm, D. K.; Petersen, L. M.; Klitgaard, A.; Knudsen, P. B.; Jarczynska, Z. D.; Nielsen, K. F.; Gotfredsen, C. H.; Larsen, T. O.; Mortensen, U. H. *Chem. Biol.* 2014, *21* (4), 519.

SUPPORTING INFORMATION

Transforming the lovastatin producing PKS, LovB, into a PKS-NRPS hybrid

Maria Lund Nielsen[1,2], Thomas Isbrandt[1,2], Uffe Hasbro Mortensen[1], Jakob Blæsbjerg Hoof[1], Mikael Rørdam Andersen[1,*], and Thomas Ostenfeld Larsen[1,*]

[1] Department of Biotechnology and Biomedicine, Technical University of Denmark, Søltofts Plads, 2800 Kongens Lyngby, Denmark

[2] Co-first author

*Corresponding authors

Contents:

Figure S1. Co-expression of LovB and LovC leads to production of dihydromonacolin L acid in A. nidulans. Extracted ion chromatograms (EICs) ([M+H]+ = 325.2373±0.001 Da) of extracts of A. nidulans expressing LovB (trace i), and LovB/LovC (trace ii). The NID3 reference strain displayed no production of dihydromonacolin L acid (trace iii).

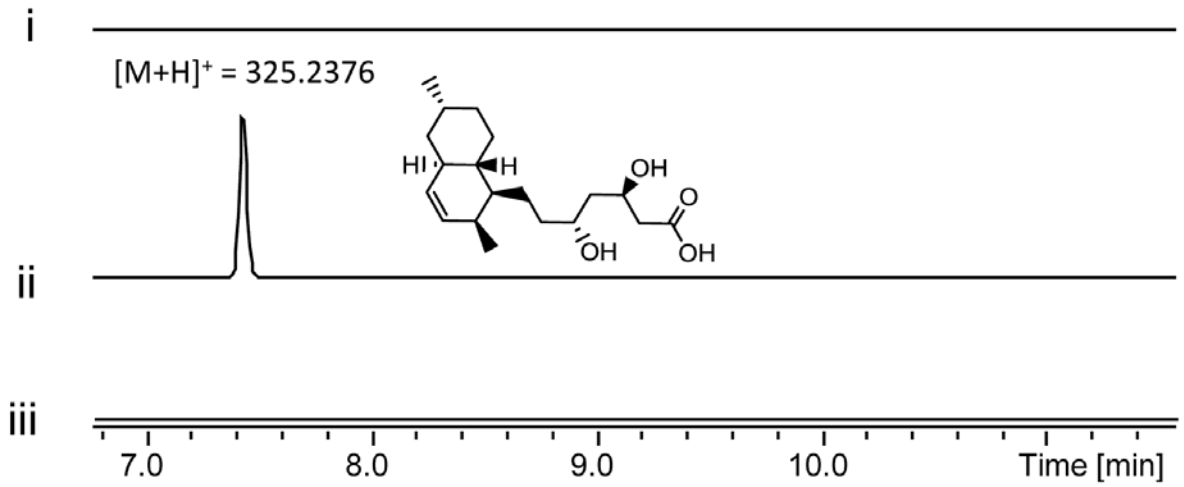
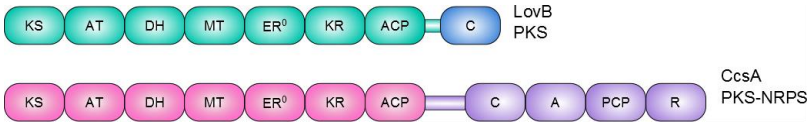
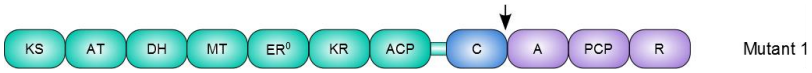


Figure S2, related to Figure 2. Design and construction of LovB-CcsA fusions. An amino acid alignment of LovB and CcsA in the regions of the transition points are shown, and the sequences of mutant 1 and 2 are indicated by a black box. A) Mutant 1 was constructed by fusing LovB to CcsA immediately downstream of the LovB endpoint. Thus, the full-length LovB including the C-terminal C domain was fused to the A-PCP-R domains of CcsA. B) Mutant 2 was constructed by replacing the native LovB C domain with the NRPS module from the CcsA PKS-NRPS hybrid. The regions between the ACP- and C domains of LovB and CcsA are highly variable, and to prevent any potential restraints of the shorter LovB linker, the linker of CcsA was preserved. The locations of the ACP- and C domain are highlighted in yellow while the predicted linker sequence of CcsA is highlighted in magenta. Amino acid alignments were constructed using Clustal Omega, while the locations of domains were predicted using CDD from NCBI.



A) Mutant 1

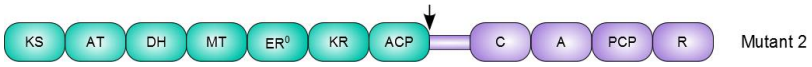



```
CcsA    NELGAAR----SATHSPLFQVLVNYRAGVSERRSFCNCDSKVLTFEQGQTPYDLSLDVID
LovB    DQLGLEVPVPTSNQPAPLLQAVFDYKQGQAESGTIGGAKITEVIATRERTPYDVVLEMSD
        ::**     *  :**:*.:*: * :*  :: .. :  : :****: *:: *

CcsA    NPGGDCHVIMAGQSVLYGAEHVAVLRGVYQNLLVAFSRNPALRLNVPPLYDTDEVKHAIK
LovB    DPTKDPLLTAKLQSSRYEAHHPQAFLESYMSLLSMFSMNPALKLAT----------
        :* * :   ** * *.* .:  * .** ** ****.*

CcsA    LGHGPAYNYQWPATIPERIDEIVERYPTHVALIDGDGRKMSYTEMARRVNTLAVVLLRQD
LovB    ------------------------------------------------------------
```

B) Mutant 2




```
                ACP domain
CcsA    VAVDIRSWFIKELQVEIPVLKILSGATVGEMVTQAQELLPKELTPNLDPNAEAKPSKPKN
LovB    GAVTVGTWFSKQLYLDLPLLKVLGGASIADLADEAAARLPPSSIPLVAATDGGAEST-DN
        ** : .** *.* ...*.**.* .**:.:. .*  ** . * : . . *. .*

CcsA    TVQPKQQTKKTIQLQNVAKAPQPALSQQVSSGVQNMIKTNPPKEAEAKQPRPEVKQAAPK
LovB    T----------------------------------------------SENEVSG--RE
        *                              . **.  :

CcsA    DSQYPTALETPSKLQDPSRNIVVAKDLAAEEKHLTDQEPVPSNMSSSSWSEIDESEGKVE
LovB    DT----DLSAAATITEPSS----------------------------ADEDDTEPG---
        *:   *.: : .:.**                  .* * . *

                                                C domain
CcsA    TSSSSSSTSASQIITKTKPVEVKKSVPMAFGQSRFWFLRHYLEDPSSFNITVSIQLDGPL
LovB    ----------------DEDVPRSHHPLSLGQEYSWRIQQGAEDPTVFNNTIGMFMKGPI
                .  :.  *.::**.  * .::  ***. ** *.: : .**:
```

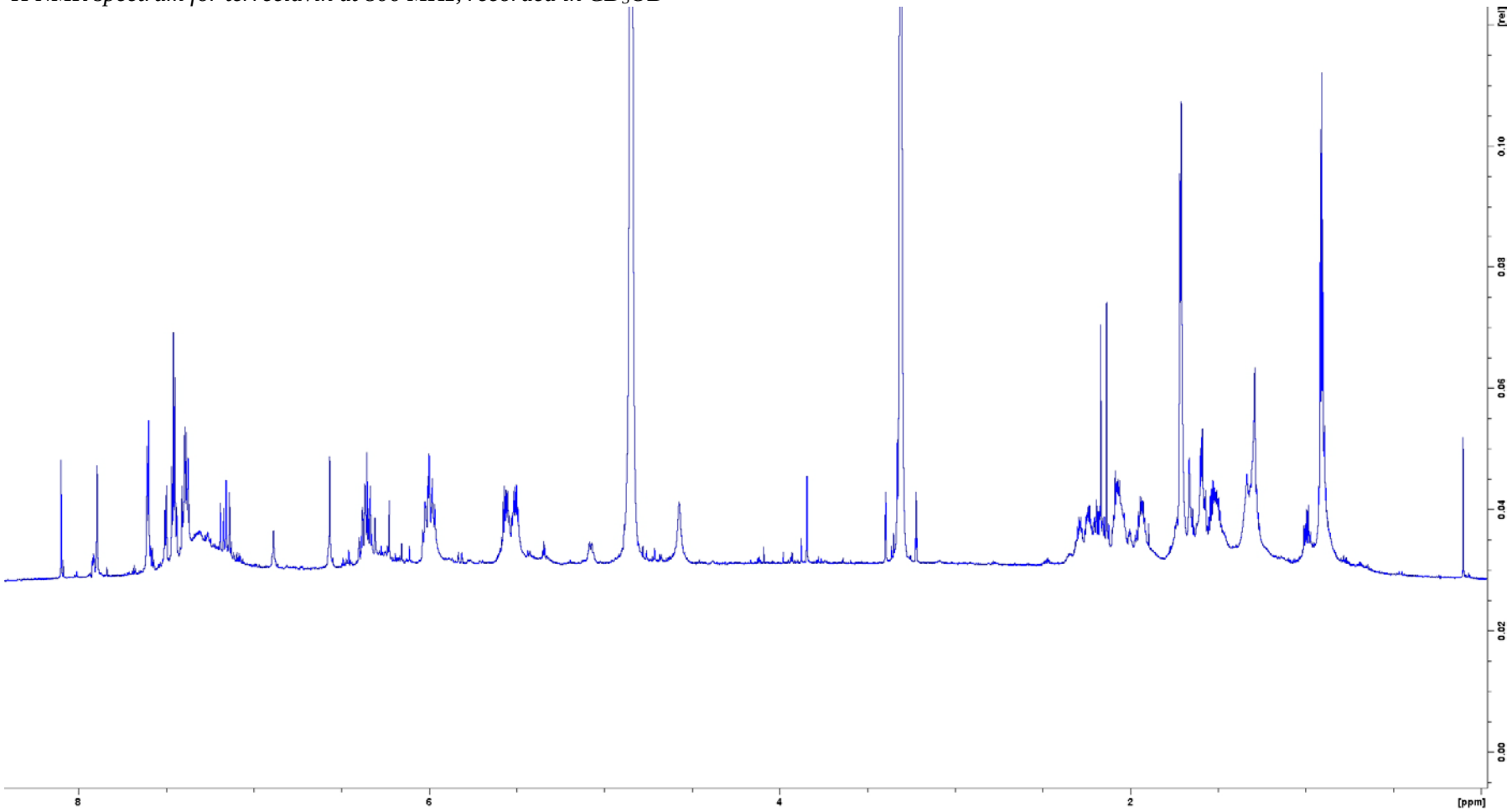
Table S1, related to Figure 4. NMR shifts and correlations. Proton and carbon shifts, as well as DQF-COSY, H2BC, HMBC and NOESY correlations for terreclavin.

#	1H	13C (HSQC)	Mult.	J [Hz]	COSY	H2BC	HMBC	NOESY
1	1.71	17.6	d	6.5	**2**	**2**	**2, 3**	**2, 3**
2	5.56	127.3	m	-	**1, 3**	**1, 3**	**1, 4**	**1, 4**
3	6.02	132.8	m	-	**2**	**2**	**1**	**1, 5**
4	5.99	133.0	m	-	**5**	**5**	**6, 2**	**1, 6a, 6b, 2**
5	5.51	130.5	m	-	**6a, 6b, 4**	**6, 4**	**7, 6, 3**	**1.55, 6a, 6b, 3**
6a	1.93	40.7	m	-	**7, 6b, 5**	**7, 5**	**7-Me, 7, 8, 5, 3**	**7-Me, 8b, 6b**
6b	2.09	40.7	m	-	**7, 6a, 5**	**5**	**7-Me, 7, 8, 5, 3**	**7-Me, 8b, 6a**
7	1.54	33.9	m	-	**7-Me, 8a**	**7-Me, 8, 6**	**7-Me, 9, 8, 6, 5**	**8a, 6a**
7-Me	0.91	19.6	m	-	**7**	**7**	**7, 8, 6**	**8a, 8b, 6a, 6b, 9a, 9b**
8a	1.29	36.3	m	-	**8b, 9a**	**9**	**7-Me, 9, 7, 6**	**7-Me, 8b, 9a, 9b**
8b	1.51	36.3	m	-	**8a, 9b**	**9**	**7-Me, 9, 7, 6**	**8a, 9a**
9a	2.24	31.6	m	-	**8a, 10**	**8**	**7, 8, 10**	**7-Me, 8a, 8b, 10**
9b	2.29	31.6	m	-	**8b, 10**	**8**	**7, 8, 10**	**7-Me, 8a, 8b, 10**
10	6.35	148.7	m	-	**9a, 9b**	**9**	**9, 12**	**9a, 9b, 12**
11	7.17	126.3	dd	26.2, 15.3	**12**	**12**	**14**	**13, 12**
12	7.40	146.3	m	-	**13, 11**	**11**	**10, 14**	**10, 11**
13	6.38	129.1	m	-	**7.4**	**12**	**11w, 12**	**11.00**
14	-	186.7	-	-	-	-	-	-
15	-	132.1	-	-	-	-	-	-
16	-	171.6	-	-	-	-	-	-
1'	7.89	146.6	s	-	-	-	**3', 18, 16, 14**	**3'**
2'	-	136.3	-	-	-	-	-	-
3'	6.57	121.9	s	-	-	-	**5', 2'w, 1'**	**5', 1'**
4'	-	135.4	-	-	-	-	-	-
5'	7.61	130.7	d	7.6	**6'**	**6'**	**3', 7'**	**3', 6'**
6'	7.46	129.9	quart	7.6	**7', 5'**	**7'**	**6', 4'**	**5'**
7'	7.38	130.3	m	-	**6'**	**6'**	**5'**	-
NH	-*	-	s	-	-	-	-	-

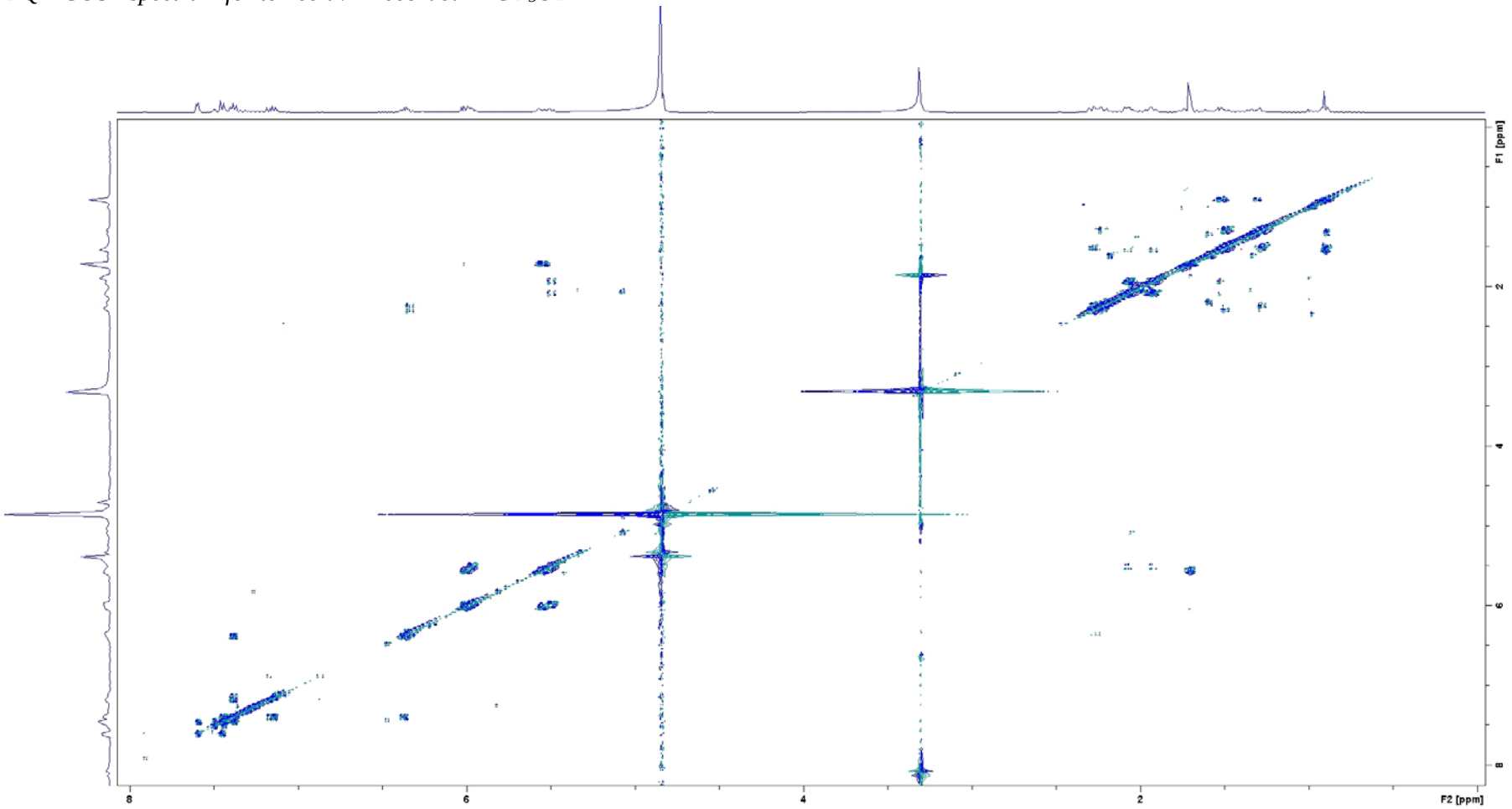
* Shift measured in DMSO-*d6*: 10.68 ppm

Table S2, NMR spectra for terreclavin

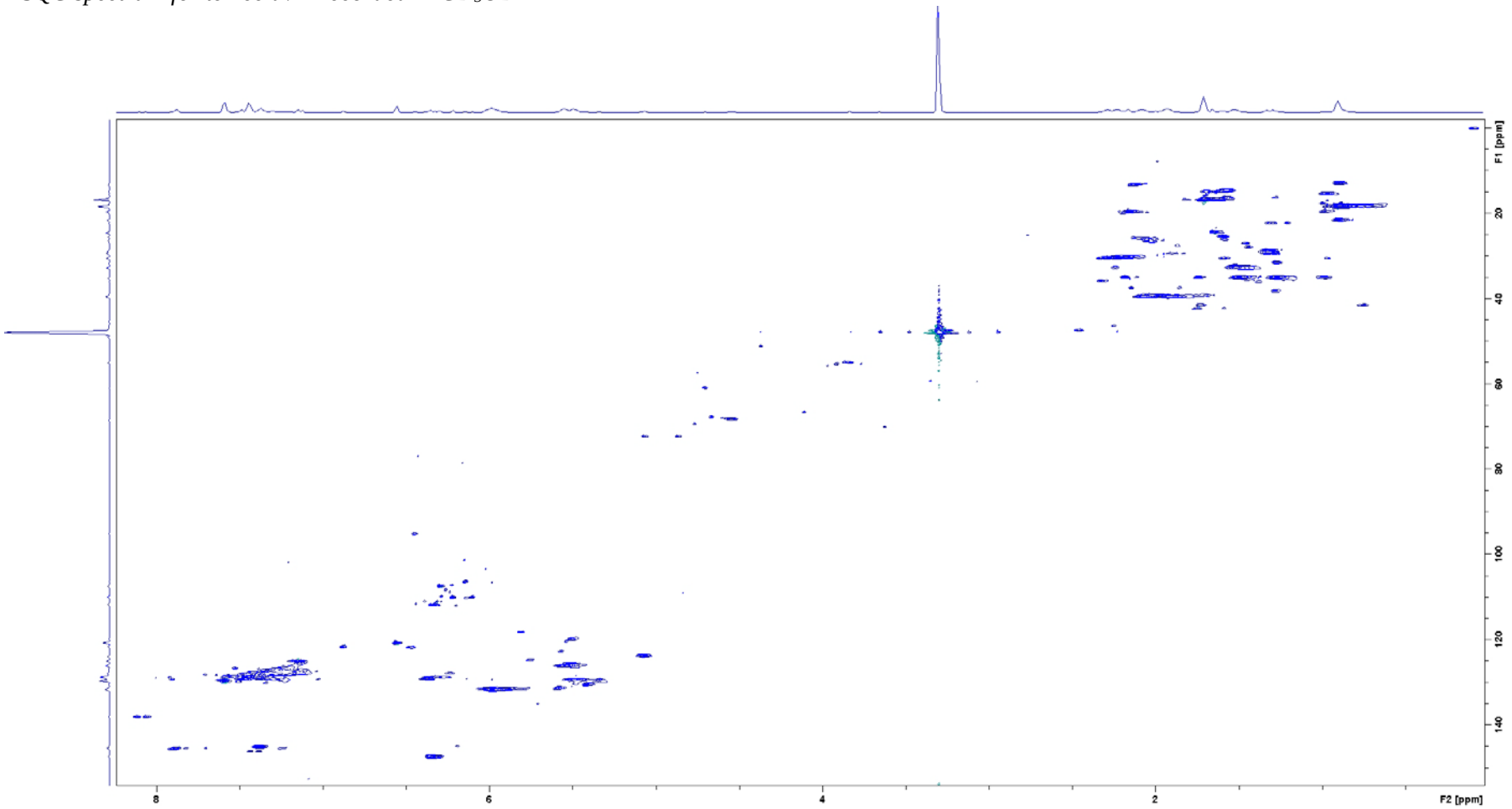
^{1}H NMR spectrum for terreclavin at 800 MHz, recorded in CD_3OD



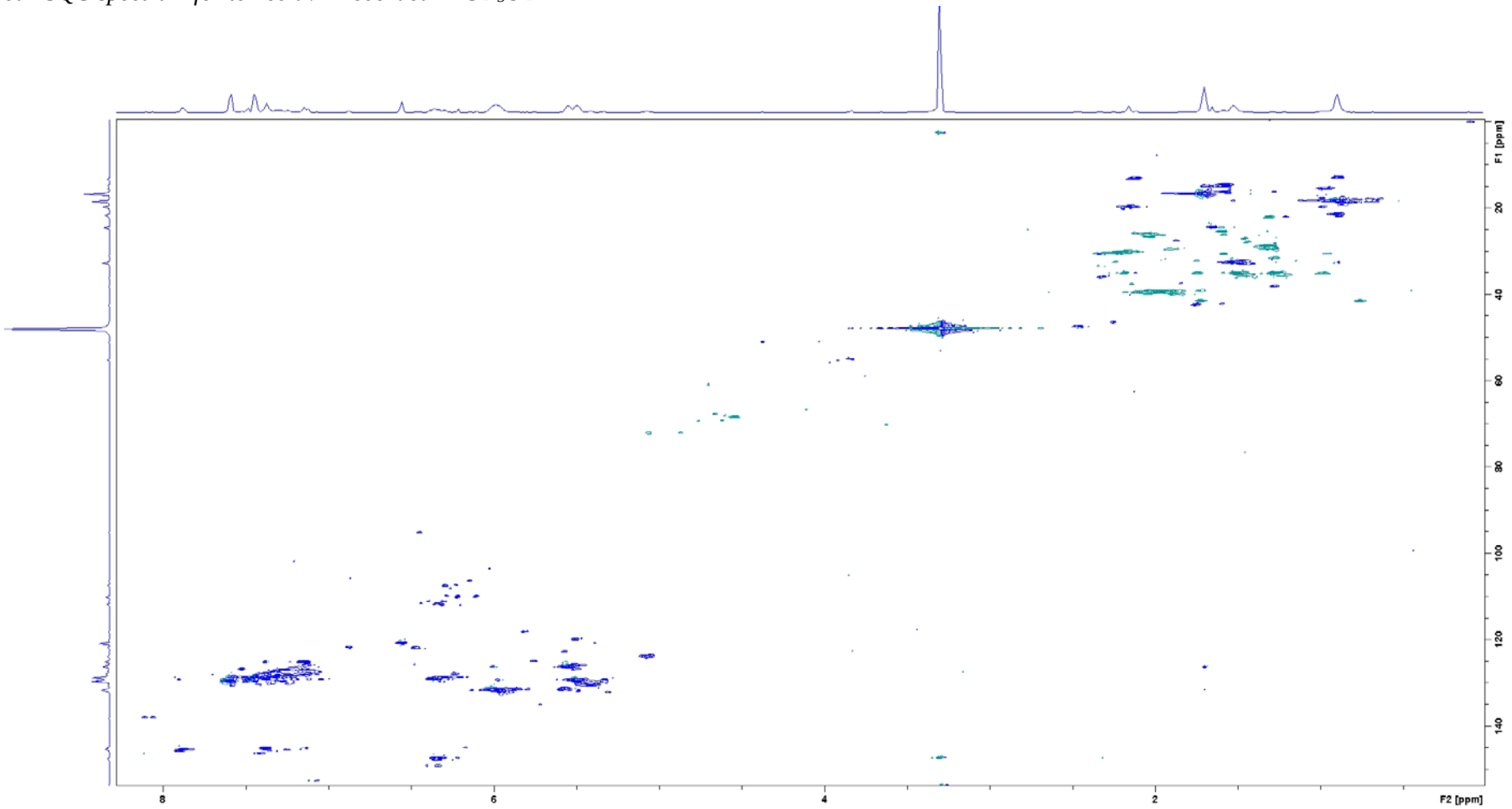
DQF-COSY spectrum for terreclavin recorded in CD_3OD



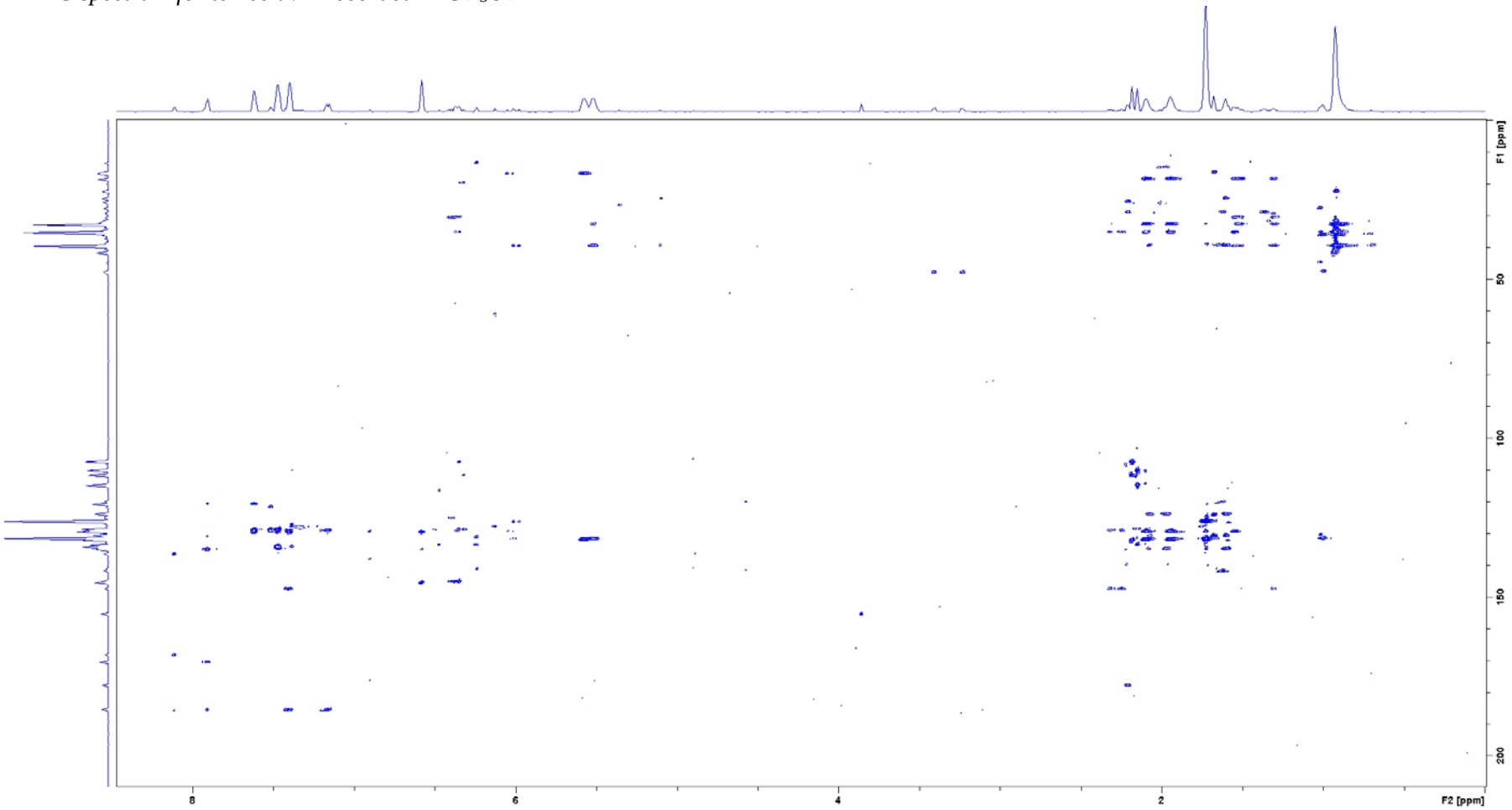
HSQC spectrum for terreclavin recorded in CD_3OD



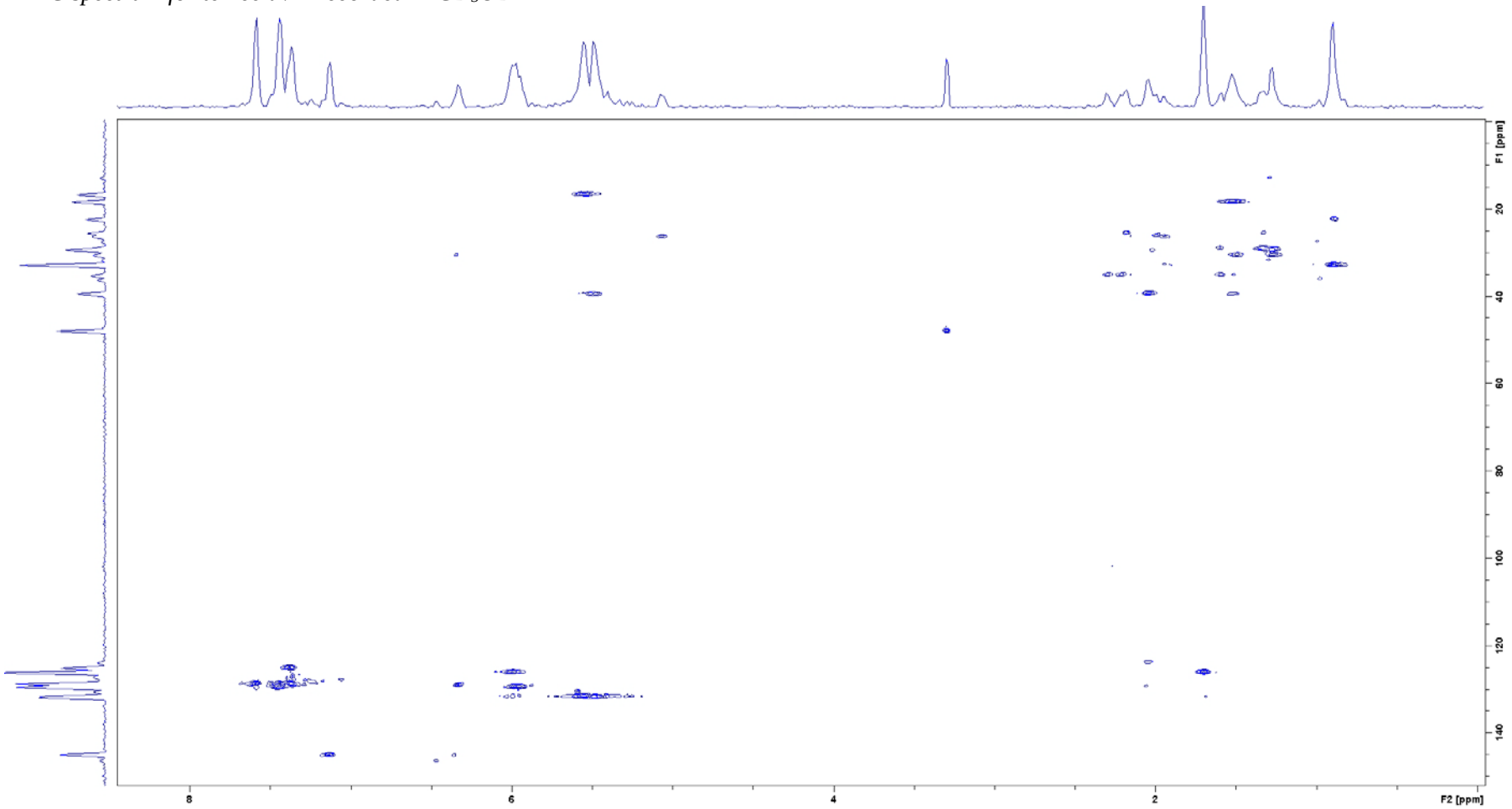
edHSQC spectrum for terreclavin recorded in CD_3OD



HMBC spectrum for terreclavin recorded in CD_3OD



H2BC spectrum for terreclavin recorded in CD_3OD



NOESY spectrum for terreclavin recorded in CD_3OD

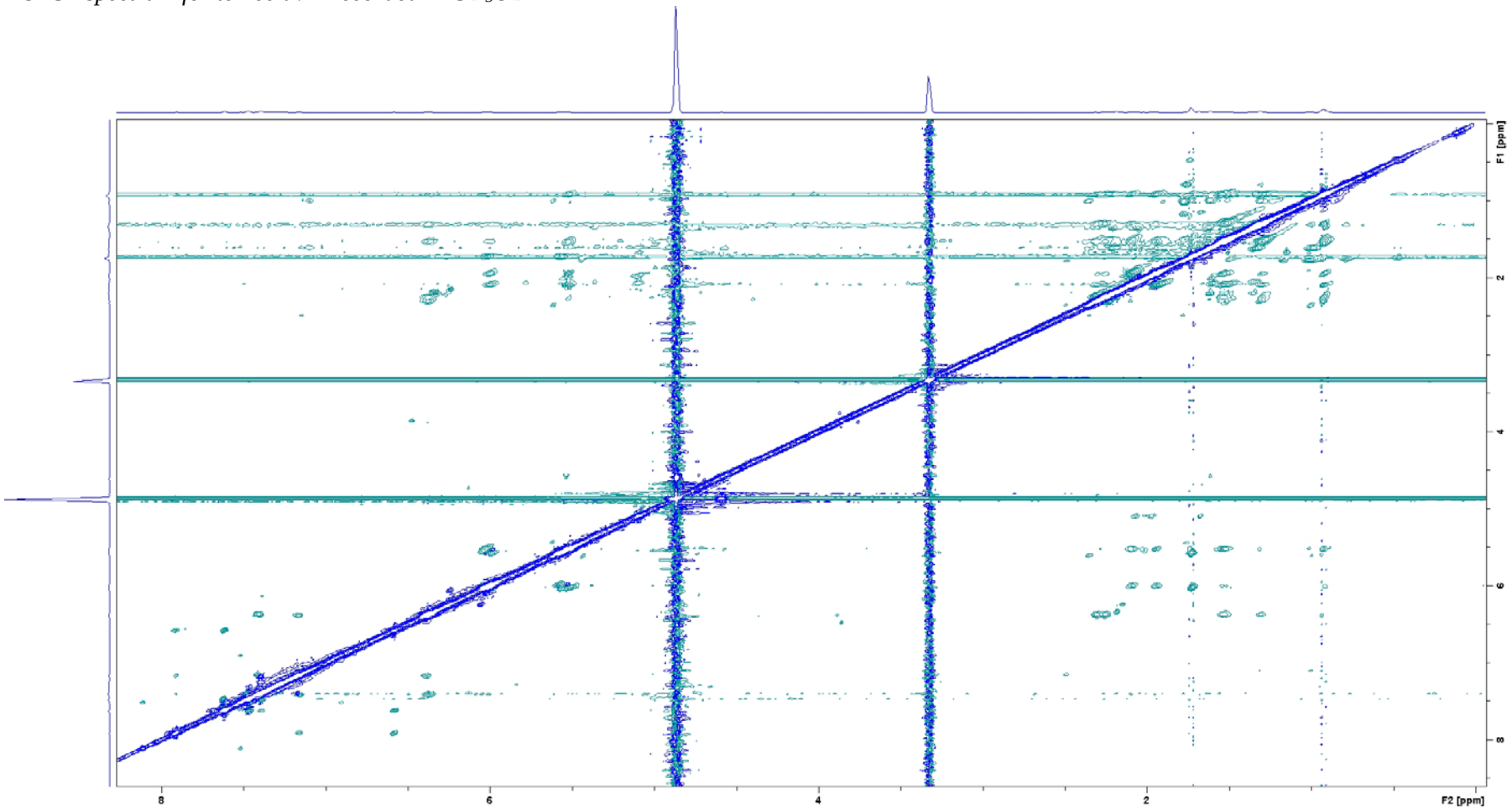


Table S3, related to Experimental Procedures. List of primers

lovC overexpression:

lovC-FU	GGGTTTAAUATGGGCGACCAGCCATTC	lovC into pU2115-2
lovC-RU	GGTCTTAAUTTTATCCCAATCCACGCAGC	lovC into pU2115-2

lovB overexpression:

lovB-Pac-FU	GGGTTTAAUATGGCTCAATCTATGTATCCTAATGAG	lovB into pU2115-5
lovB-Pac-RU	GGTCTTAAUCTCTTCGTTCGGCTCACCATA	lovB into pU2115-5
lovB-2FU	ACGGAGTTCUATACCAACACACTCA	lovB into pU2115-5
lovB-2RU	AGAACTCCGUCAGGAGCCCA	lovB into pU2115-5

Mutant 1 overexpression:

lovB-Pac-FU	GGGTTTAAUATGGCTCAATCTATGTATCCTAATGAG	lovB into pU2115-5
ccsA-Pac-RU	GGTCTTAAUTGCTGTGTCCCAATCAGACGT	Construction of Mutant 1 and 2
lovB-4RU	ACTGCCAGCTUCAGGGCG	lovB-ccsA chimera overexpression
ccsA-14FU	AAGCTGGCAGUCCCGCCCCTATATGATACC	lovB-ccsA chimera overexpression

Mutant 2 overexpression:

lovB-Pac-FU	GGGTTTAAUATGGCTCAATCTATGTATCCTAATGAG	lovB into pU2115-5
ccsA-Pac-RU	GGTCTTAAUTGCTGTGTCCCAATCAGACGT	Construction of Mutant 1 and 2
lovB-3RU	AGCGGTGGCUGCGACGAGGG	Construction of Mutant 2
ccsA-13FU	AGCCACCGCUGAGGCAAAGCCGTCG	Construction of Mutant 2

Table S4, related to Figure 2. Lists of plasmids and strains

Plasmids

Plasmid #	Plasmid name	Purpose
pAC286	pU2115-2	Cloning vector
pAC290	pU2115-5	Cloning vector
pAC507	pU2110-2-lovC	Overexpression of *lovC*
pAC501	pU2110-5-lovB	Overexpression of *lovB*
pAC505	pU2110-5-LC2 (mutant 1)	Overexpression of mutant 1
pAC503	pU2110-5-LC1 (mutant 2)	Overexpression of mutant 2

Strains

Strain #	Strain identifier	Genotype
1	nkuAΔ (NID1)	*argB2*, *pyrG89*, *veA1*, *nkuA*Δ
2	nkuA-trS	*argB2*, *pyrG89*, *veA1*, *nkuA-trS::AFpyrG*
3	lovC prepop	*argB2, pyrG89, veA1, nkuA*Δ*, IS2::PgpdA-lovC-TtrpC::AFpyrG*
4	lovC pop	*argB2, pyrG89, veA1, nkuA*Δ*, IS2::PgpdA-lovC-TtrpC*
5	lovB prepop	*argB2, pyrG89, veA1, nkuA*Δ*, IS2::PgpdA-lovB-TtrpC::AFpyrG*
6	lovB, lovC prepop	*argB2, pyrG89, veA1, nkuA*Δ*, IS5::PgpdA-lovB-TtrpC::AFpyrG*, *IS2::PgpdA-lovC-TtrpC*
7	LC1, lovC prepop	*argB2, pyrG89, veA1, nkuA*Δ*, IS5::PgpdA-LC1-TtrpC::AFpyrG*, *IS2::PgpdA-lovC-TtrpC*
8	LC2, lovC prepop	*argB2, pyrG89, veA1, nkuA*Δ*, IS5::PgpdA-LC2-TtrpC::AFpyrG*, *IS2::PgpdA-lovC-TtrpC*

Strain #	Strain notes
1	Permanent nku deletion strain (background strain)
2	Transient small repeat in nkuA (reference strain for chemical analysis)
3	Oex of *A. terreus lovC*
4	Oex of *A. terreus lovC*
5	Oex of *A. terreus lovB*
6	Oex of *A. terreus lovC* and *lovB*
7	Oex of *A. terreus lovC* and *lovB-ccsA* chimera (mutant 2)
8	Oex of *A. terreus lovC* and *lovB-ccsA* chimera (mutant 1)